

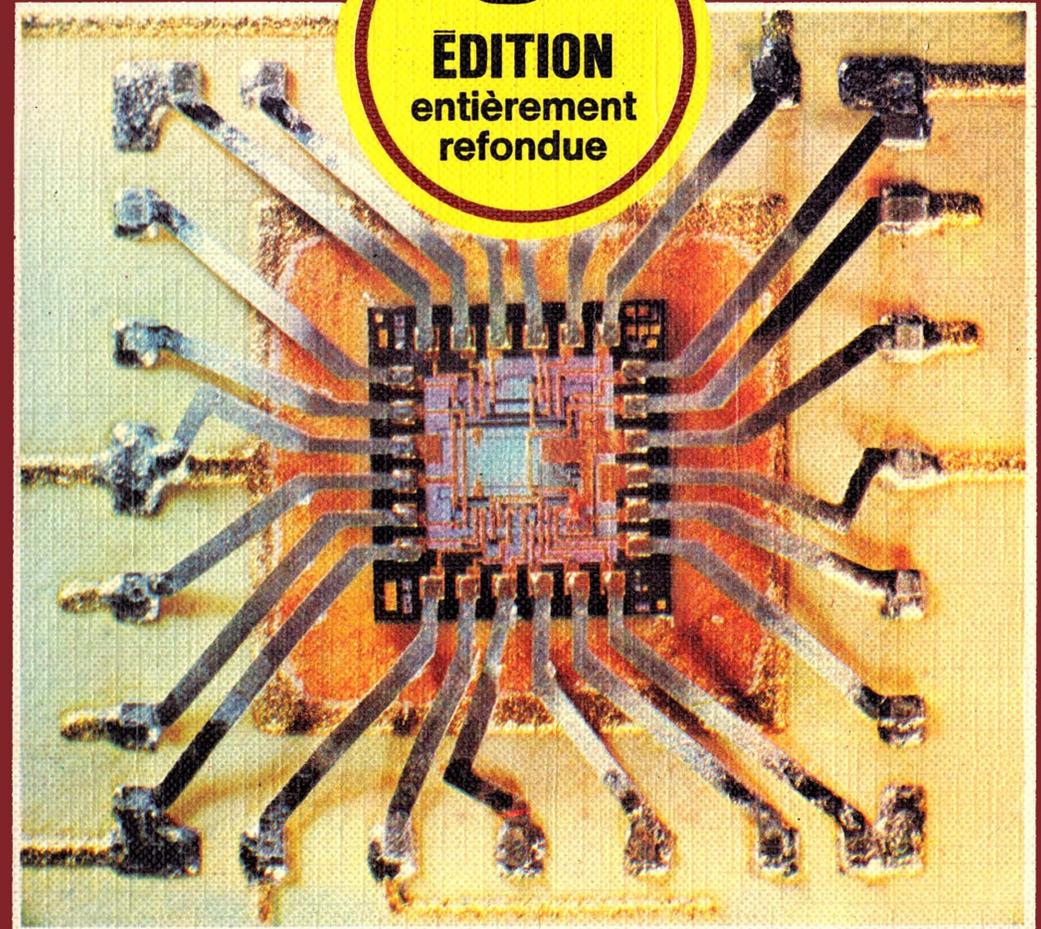
H. LILEN

circuits intégrés JFET-MOS-CMOS

PRINCIPES ET APPLICATIONS

3^e

ÉDITION
entièrement
refondue



Tout ce qu'il faut savoir des circuits intégrés à transistors à effet de champ, JFET, MOS et MOS complémentaires :
— caractéristiques, intégration, logiques résultantes ;
— applications à la logique, aux mémoires, aux microprocesseurs et au domaine linéaire, etc.

Ce livre traite également des VMOS, des circuits à transfert de charges et des toutes dernières technologies.

ISBN 2 7091 0760 0

JFET-MOS-CMOS

circuits intégrés


EDITIONS
RADIO


EDITIONS RADIO

circuits intégrés
JFET-MOS-CMOS
PRINCIPES ET APPLICATIONS

AUTRES OUVRAGES DU MÊME AUTEUR

- Principes et applications des circuits intégrés linéaires.
- Circuits intégrés numériques.
- Circuits hybrides.
- Du microprocesseur au micro-ordinateur.
- Guide mondial des microprocesseurs.
- Mémoires intégrées.
- Programmation des microprocesseurs.

en collaboration avec R. Arouete :

- Théorie et pratique des microprocesseurs.

En couverture de ce livre :

L'équipe parisienne d'Honeywell-Bull (à Saint-Ouen), a mis au point un système de traitement automatique des puces de circuits intégrés qui les prend totalement en charge, de la tranche initiale de silicium, au montage définitif sur circuit hybride. Ce « micro-packaging » se sert d'un film du type cinéma comme support intermédiaire.

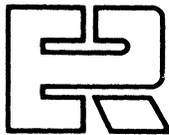
La photo de couverture représente ainsi une puce de CI qui, après avoir été « manipulée » sans aucune intervention humaine dès sa complète élaboration sur la pastille de silicium, a été reportée toujours automatiquement sur un circuit hybride et soudée.

H. LILEN

circuits intégrés
JFET-MOS-CMOS

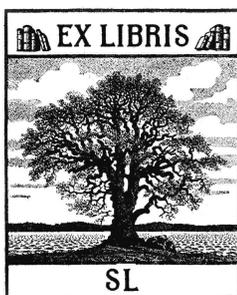
PRINCIPES ET APPLICATIONS

3^e édition



Editions Radio

9, RUE JACOB - 75006 PARIS



© Éditions Radio, 1979

Imprimé en France
Imprimerie Berger-Levrault, Nancy

Dépôt légal : 1^{er} trimestre 1979
Éditeur n° 760 - Imprimeur n° 778135
I.S.B.N. 2 7091 0760 0

LISTE DES CHAPITRES

Avant-propos	7
1^{re} partie — FET et circuits intégrés	
I. Introduction aux FET	9
II. Fonctionnement du MOS fondamental, canal P	27
III. Les divers types de MOS	51
IV. La fabrication des circuits intégrés	68
V. L'arsenal des technologies	89
VI. Les JFET et leurs familles	121
VII. Circuits à transfert de charge : CCD, BBD et variantes	141
VIII. Les VMOS	167
2^e partie — La logique intégrée	
IX. L'inverseur fondamental	179
X. Les portes logiques	211
XI. Les flip-flops	235
XII. Les registres	263
XIII. Les MOS complémentaires et la logique	283
3^e partie — Les applications	
XIV. Les mémoires mortes	327
XV. Les mémoires vives	345
XVI. Les microprocesseurs	359
XVII. Télécommunications : les CODEC	379
XVIII. Applications linéaires	393
Bibliographie	409
Table des matières détaillée	411

AVANT-PROPOS

Ce volume, qui paraît en 1979, a succédé à un livre plus ancien dont la première édition date de 1972 et dont le titre était « Circuits intégrés MOS et CMOS ».

La fin des années 60 et les premières années 70 ont, on l'a constaté avec le recul du temps, constitué l'époque héroïque des MOS. Puis ceux-ci se sont affirmés à tel point qu'ils ont conquis, sinon créé, d'immenses marchés.

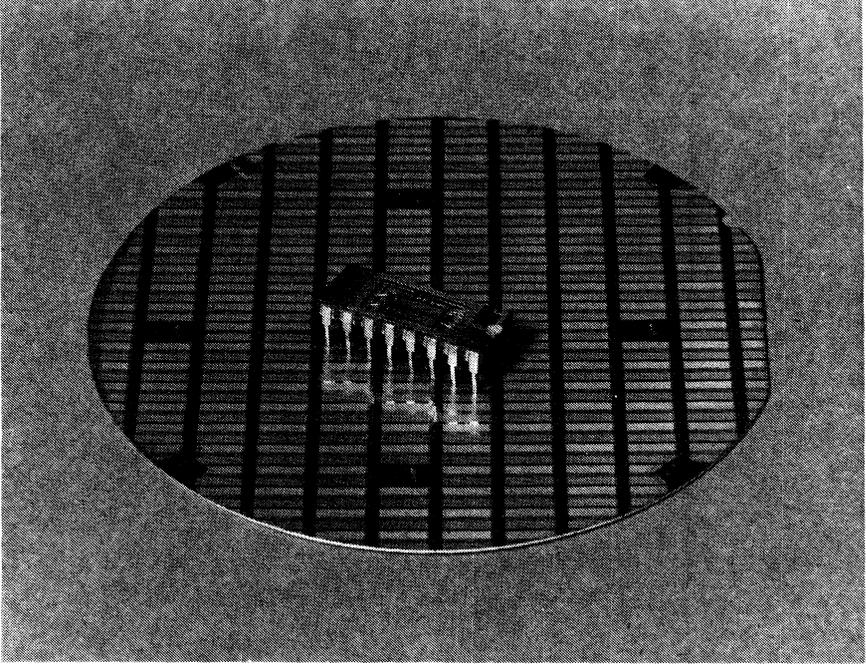
Voici quelques années, les PMOS dominaient : aujourd'hui, leur courbe de maturité a abordé l'autre versant et ils cèdent le pas aux NMOS. Les MOS complémentaires tentaient de partir à l'attaque des positions TTL car on n'avait pas encore compris qu'ils peuvent aussi autre chose et mieux. Les circuits à transfert de charge n'étaient que curiosité de laboratoire. Nul ne songeait encore aux VMOS, encore moins d'ailleurs aux BIFET ou BIMOS qui mêlent hardiment des technologies unipolaires et bipolaires pour des applications linéaires, cette fois. Les premiers microprocesseurs apparaissaient tout juste...

L'évolution a donc été extrêmement rapide. Aussi, plutôt que de publier une 3^e édition des « Circuits intégrés MOS et CMOS » a-t-on jugé préférable de remanier cet ouvrage pour en présenter une version nouvelle et à jour, sous un titre plus général.

L'accent a été mis sur les technologies les plus récentes et leurs applications, par exemple sur les VMOS sous leur double aspect, d'ailleurs, intégré et discret. Les applications aux fonctions linéaires ont été développées via les structures BIMOS et BIFET; etc. Par contre, certains chapitres ont été allégés, sur les mémoires par exemple; en effet, les circuits mémoires ont pris une telle importance qu'il a fallu leur consacrer un volume spécial (aux Éditions Radio); aussi, les lecteurs qui souhaiteraient aller au-delà de la synthèse qui en est présentée ici pourront s'y reporter. Il en ira de même, d'ailleurs, pour les microprocesseurs.

A travers l'actualisation de ce livre, on peut juger de la cadence à laquelle évolue l'électronique. Quand on songe qu'on ne voit aucune raison pour que ce rythme se ralentisse...

L'AUTEUR



Le record en 1978 : la mémoire RAM de 64 kilobits de Texas, vue ici sur sa tranche initiale de silicium.

PREMIÈRE PARTIE

FET ET CIRCUITS INTÉGRÉS

CHAPITRE I

INTRODUCTION AUX FET

Ce chapitre sert à présenter, en synthèse, les transistors à effet de champ (FET) dans leurs variantes essentielles, ainsi que leurs caractéristiques générales et leur histoire.

1. LES TRANSISTORS UNIPOLAIRES

Les transistors qu'utilise l'électronique sont de deux types :

1. *Les bipolaires*, les plus anciens commercialisés, sont les PNP et NPN classiques. Dans les bipolaires, la conduction s'explique à l'aide de porteurs des deux types, N et P, d'où leur nom.

2. *Les unipolaires*, dont l'origine est bien plus ancienne mais qui n'ont été commercialisés qu'après les bipolaires. Dans les unipolaires, la conduction fait appel à un seul type de porteur de charge, N ou P. Ils sont regroupés sous l'appellation générale de *transistors à effet de champ*, ou TEC, mais plus fréquemment FET à partir de leur désignation américaine de « *Field Effect Transistor* ».

Le FET est donc un composant actif qui existe aussi bien en versions discrètes qu'intégrées. Dans ce dernier cas, sa fabrication relève des processus habituels de l'intégration. A partir du procédé Planar, cette fabrication est résumée dans la figure 1; une tranche initiale de silicium de 7,5 ou 10 cm de diamètre et de 1/3 mm d'épaisseur sert de support; par des procédés photolithographiques sont définies les zones qui seront dopées; des croissances d'oxyde et des dépôts d'aluminium complètent ce processus. Enfin, la surface de la tranche silicium est divisée en pavés (ou *puces*) qui représentent chacun un circuit intégré; elle est découpée et chaque puce est fixée sur un support puis dans un boîtier.

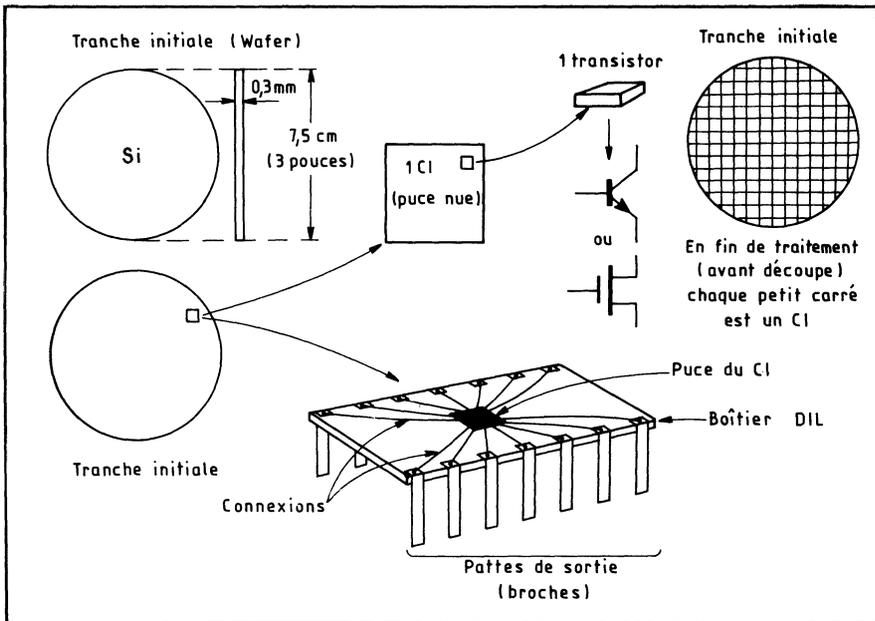


Fig. 1. — De la tranche initiale de silicium au circuit intégré.

Si l'on examine une coupe de ce silicium en un point précis et sur quelques dizaines de microns de large, on retrouve la structure d'un transistor de base, stylisée dans la figure 2 pour un NPN bipolaire, dans la figure 3 pour un JFET, un MOS classique et un FET Schottky, et dans la figure 4 pour un seul MOS, à canal N. C'est qu'en effet, les FET se classent à leur tour en plusieurs catégories :

1. *FET à jonction* : l'électrode de commande est en contact avec le semiconducteur, le silicium. L'impédance d'entrée est donc celle d'une diode, mais polarisée en inverse. On désigne les FET à jonction par le sigle JFET, ou encore parfois JUGFET, pour « *junction gate FET* ».

2. *FET à électrode de commande isolée*, le plus souvent par de la silice, du semiconducteur. Ce sont les IGFET, pour « *insulated gate field effect transistor* »,

Fig. 2. — Le transistor bipolaire intégré.

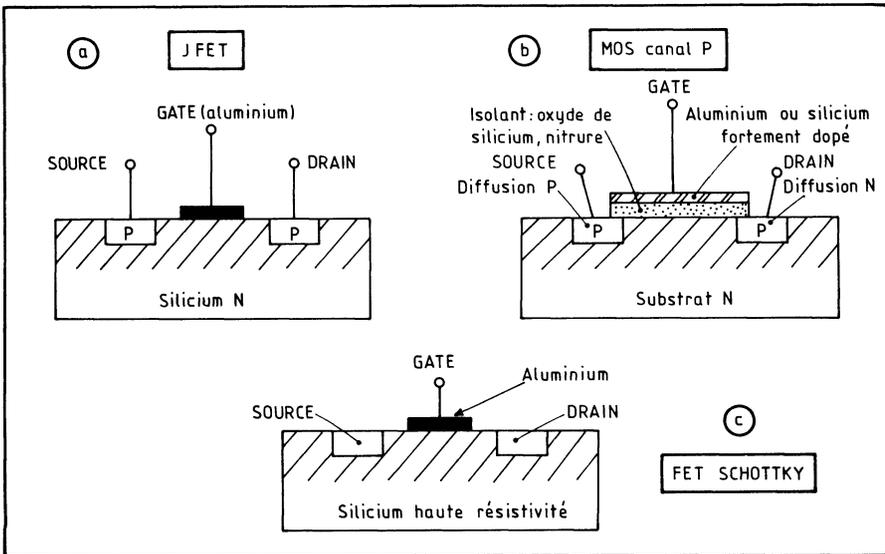
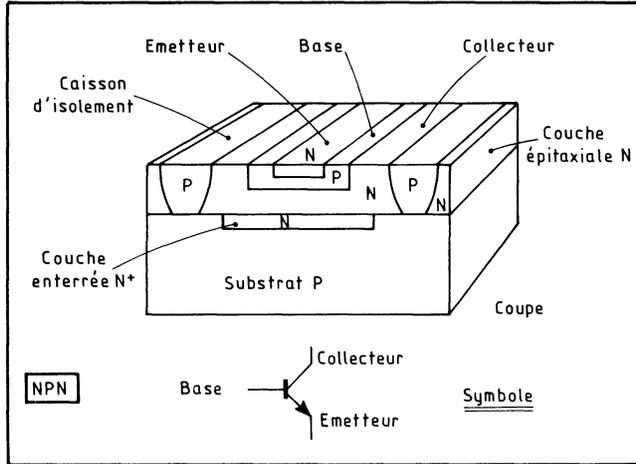
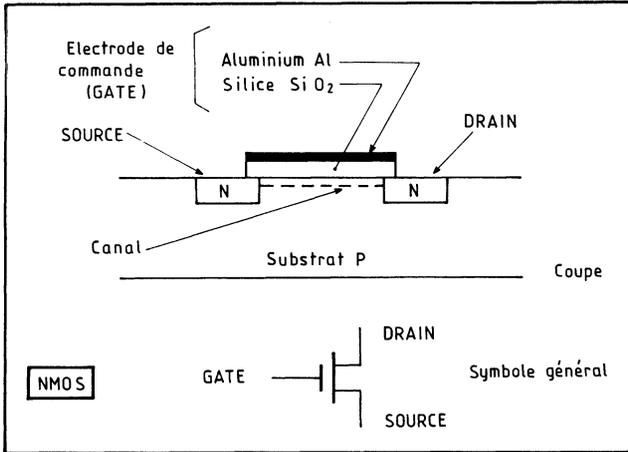


Fig. 3. — Le MOS canal P comparé au FET à jonction et Schottky.

soit *FET à gate isolé*. L'impédance d'entrée est la forte résistance de l'isolant avec, en parallèle, une capacité.

En sous-catégories, on trouvera (ce qui est résumé par la figure 5) *pour les FET* :

- *Le JFET Schottky, ou MESFET* : la diode d'entrée est une Schottky.
- *Le VFET ou Gridistor*, type de FET à grille de commande dans le silicium. *et pour les IGFET* :



4. — Principe du MOS.

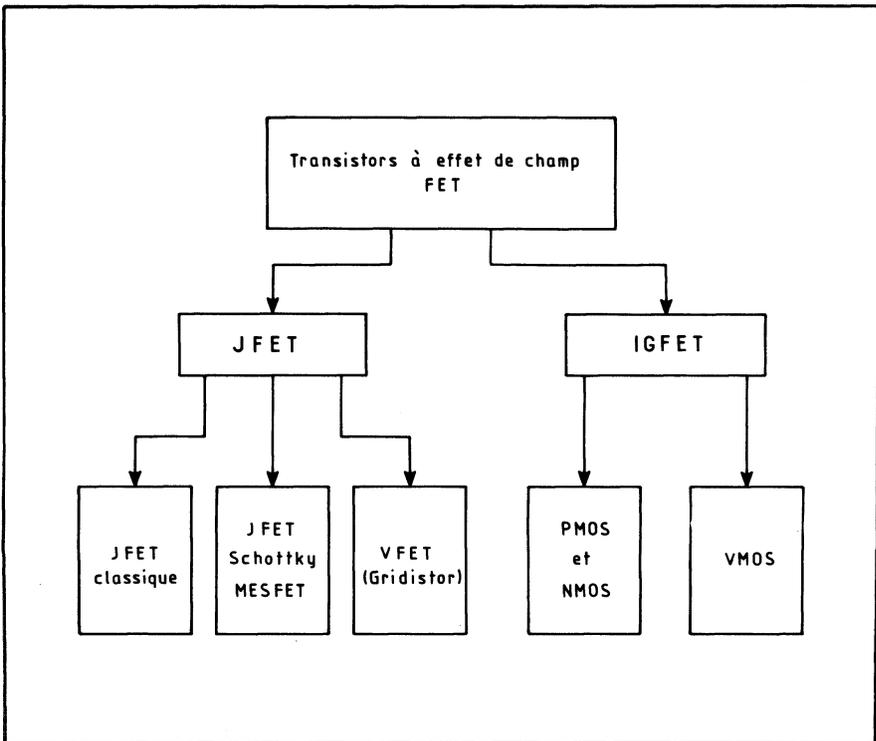


Fig. 5. — Le classement des transistors à effet de champ.

- *Le transistor MOS*, le sigle MOS définissant sa structure de base : *métal-oxyde-silicium*. On y ajoute parfois : *transistor*, ce qui donne le sigle MOST, ou même MOSFET si l'on tient à préciser qu'il s'agit d'un FET, ce qui est redondant. C'est le composant de loin le plus répandu, à ce point que son nom désigne même ses variantes qui, en toute bonne logique, ne répondent plus à cette définition.

- *Le VMOS* : c'est le « *MOS Vertical* », commercialisé à partir de 1976.

Comme on a pu le constater sur les schémas de structure, les FET sont des transistors à trois électrodes appelées *source*, *drain*, et pour l'électrode de commande, *gate*. On a conservé ici cette désignation américaine qui signifie *porte* pour une double raison : le mot *grille* qu'on lui substitue parfois n'est pas physiquement conforme à sa structure (sauf pour les Gridistors), alors qu'il s'applique parfaitement à l'électrode de commande de tubes; d'autre part, le mot *porte* a le défaut de ne pas commencer par la même initiale, ce qui risque de créer des confusions dès que l'on passe aux formules, normalisées par l'usage.

Une autre confusion provient de l'évolution même des MOS. Alors que les premiers éléments étaient réellement constitués, vus en coupe, par *un métal* (aluminium), *un oxyde* (silice) et *un semiconducteur* (silicium), ils ont pu donner naissance à des variantes dans lesquelles l'oxyde était remplacé par du *nitru de silicium*, et l'aluminium par du *silicium dopé*. Bien qu'on leur ait conservé l'appellation de MOS, des désignations plus conformes sont :

- MIS, pour : *métal-isolant-semiconducteur*, ou encore :
- CIS, plus général, pour *conducteur-isolant-semiconducteur*.

On a vu que le canal des FET, et plus précisément pour ce qui va suivre des MOS, définit leur désignation. Or, on trouvera encore les deux importantes catégories qui sont :

- *Les MOS monocanaux*, qui regroupent soit des NMOS, soit des PMOS pour constituer des portes logiques de base (mais ou un type, ou l'autre *exclusivement*).

- *Les MOS complémentaires*, ou CMOS, dans lesquels un NMOS plus un PMOS participent à la constitution de la porte élémentaire.

2. LES AVANTAGES DES FET

Dans tous les FET, la conduction s'effectue via *un canal* qui relie la source au drain et qui est de même polarité que ces électrodes : soit P, soit N. Selon le cas, on traitera donc de *FET à canal P* ou de *FET à canal N*, et pour les MOS, de NMOS ou PMOS (la première lettre traduisant donc la polarité du canal conducteur).

L'action du gate peut viser :

- *Soit à créer ce canal conducteur* : il commande alors l'enrichissement de la zone appelée canal en porteurs. On dit qu'il s'agit d'un *FET à enrichissement*, bloqué en l'absence de signal de commande.

- *Soit à supprimer ce canal*, pré-existant, en l'appauvrissant en porteurs. On dit qu'il s'agit d'un *FET à appauvrissement*, ou d'un *FET à déplétion*, ce qui est synonyme. Il est conducteur en l'absence de signal de commande.

L'impédance d'entrée des FET est très élevée et peut dépasser le téraohm ($10^{12} \Omega$). De ce fait, on considère généralement qu'ils sont commandés par une tension, et non plus par un courant comme avec les bipolaires. D'autre part, la puissance de commande est insignifiante, toutes proportions gardées. Parmi les caractéristiques qui ont contribué à faire leur succès, plus particulièrement avec

les MOS, il convient de citer (en notant que ces comparaisons s'appliquent aux familles bipolaires TTL et ECL, mais non à la technologie I2L qui fait exception) :

1. *Leur structure est bien plus simple que celle des bipolaires.* Leur processus de fabrication est alors plus simple et ils occupent moins de surface de silicium.
2. *Leur consommation est généralement bien inférieure à celle des bipolaires.*
3. Par conséquent, *leur densité d'intégration est bien supérieure*, de 5 à 10 fois.
4. Autre conséquence, *ils sont plus économiques* (prix de revient moindre).

On ajoutera encore qu'ils autorisent des modes de fonctionnement inconnus de la TTL ou de l'ECL, le *mode dynamique*, et qu'ils se prêtent à des variantes telles que les *circuits à couplage de charge* dont il n'existe aucun équivalent en discret. En outre, les CMOS (MOS complémentaires) constituent la famille la plus économe en puissance consommée (toujours I2L exclue), et celle qui procure la plus haute immunité au bruit (cette fois, incomparablement supérieure à celle de la I2L, mais toutefois inférieure à celle des familles DTL Zener et leurs variantes).

Le grand reproche que l'on a pu adresser aux MOS a été, pendant « longtemps », leur faible vitesse. Or, dès 1977-1978 certaines familles de NMOS et de CMOS sur saphir (SOS) sont venues concurrencer les bipolaires dans des *versions commerciales* (donc disponibles et non plus en échantillons de laboratoire).

Théoriquement, certaines structures à effet de champ peuvent accéder à des fréquences de fonctionnement très élevées, ce qui s'est confirmé en laboratoire, par exemple 20 GHz (vingt gigahertz) avec des produits gain-bande de 50 GHz. Il est vrai qu'en vitesse, la suprématie appartenait toujours, en 1978, à la TTL/S et surtout à l'ECL, cette dernière devant conserver son énorme avantage encore pour de longues années puisqu'elle a alors été choisie pour les futures familles d'ordinateurs qui seront commercialisées au-delà de 1980.

La puissance a, elle aussi, été gagnée par les MOS dès 1977 avec l'apparition des VMOS, et aussi avec les VFET de S. TETZNER. En version discrète, les VMOS commandaient près de 100 W en 1978.

Les FET sont donc encore loin d'avoir dit leur dernier mot. Autour des années 1978, les recherches portaient souvent sur un accroissement général des performances, en particulier à travers les concepts ou technologies HMOS et VMOS. Puis, sur l'association FET-bipolaire afin d'obtenir, pour commencer, des amplificateurs opérationnels (donc des fonctions linéaires) tout à fait remarquables (technologies BIFET et BIMOS). Également sur l'exploitation de familles nouvelles, à MESFET par exemple, ou de nouveaux matériaux ou procédés de fabrication (le saphir des SOS); ou encore de structures dérivées (CCD).

3. CIRCUITS A TRANSFERT DE CHARGES

Bien que constituant une catégorie à part, les circuits à transfert de charge (ou CTD, de *charge transfer devices*) peuvent être considérés comme dérivant des MOS dont ils rappellent ou possèdent la structure (fig. 6). Ils comprennent :

- les CCD (de *charge coupled devices*), ou circuits à couplage de charge,
- les BBD (de *bucket brigade devices*), qu'on traduit par « chaînes à seaux » ou encore par « éléments à chapelet »,
- les SCT (de *surface charge transistor*).

Dans les CTD, on agit sur des charges électriques piégées dans le substrat. Ces circuits n'ont aucun équivalent en composants discrets et ils constituent la première création entièrement originale de l'intégration monolithique.

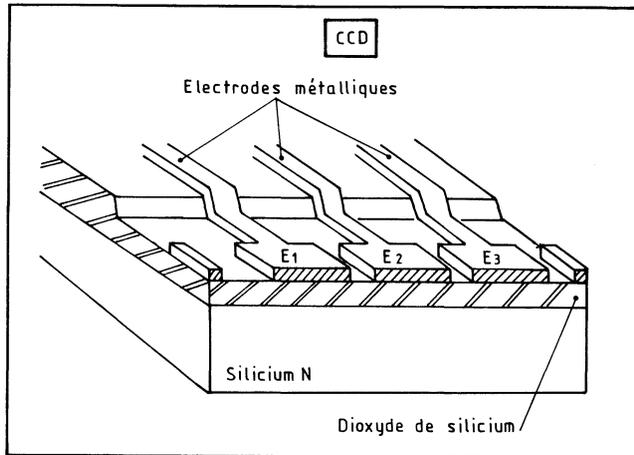


Fig. 6. — Principe des CCD.

4. SSI, MSI, LSI ET VLSI

Les circuits intégrés sont également classés selon leur complexité :

- *Le SSI* est l'intégration de fonctions simples et signifie « *single size integration* ». Par exemple, un SSI comprendra jusqu'à 25 portes intégrées sur une puce;
- *Le MSI*, ou « *medium size integration* » est l'étape au-dessus, et selon une classification maintenant assez répandue, comprendra des CI qui regroupent de 25 à 250 portes élémentaires;
- *Le LSI* est l'étape supérieure, car c'est « l'intégration à grande échelle » (*large scale integration*) qui traite de circuits de 250 à 2500 portes.
- Enfin, on a proposé le VLSI (*Very large scale integration*) pour les circuits comprenant plus de 2 500 portes.

Tableau A. — Un classement linéaire et décimal de la complexité des circuits intégrés

DEGRÉ D'INTÉGRATION	SIÈGLE	DÉFINITION	NOMBRE DE PORTES	NOMBRE DE TRANSISTORS
Simple	SSI	Single size integration	1 à 25	< 100
Moyenne	MSI	Medium scale integration	25 à 250	100 à 1000
Grande	LSI	Large scale integration	250 à 2500	1000 à 10 000
Très grande	VLSI	Very large scale integration	> 2500	> 10 000

5. LES FONCTIONS LOGIQUES FONDAMENTALES

Les structures logiques que l'on réalise avec les transistors sont :

- *Tout d'abord, les portes fondamentales* auxquelles on conservera leur appellation américaine par souci de simplicité, la plus simple étant l'inverseur, la porte AND (le *ET* logique), la porte OR (*OU* logique), puis les NAND (*non-ET*) et NOR (*non-OU*).

- *Des fonctions plus complexes* qui sont les OU exclusif, ou XOR, la porte à coïncidence, ou comparateur souvent notée COIN, etc.

- *Les bascules, ou flip-flops (FF)* : ils sont de quatre types fondamentaux : RS, JK, D et T, en variantes de base ou maître-esclave.

- *Enfin, des fonctions complexes* dérivant de ces structures.

Elles seront développées dans cet ouvrage où l'on se servira des symboles proposés par l'*American Standards Association*, parce qu'ils ont été adoptés dans les faits par les électroniciens. Ces symboles sont résumés dans la figure 7, et comparés à ceux des autres normes figure 8.

Outre ces fonctions « classiques » figurent les fonctions-mémoires, dans une palette de structures qui atteint un degré de développement exceptionnel avec les MOS : registres à décalage, RAM statiques et dynamiques, ROM, PROM et REPROM, RMM et MNOS, CCD et BBD..., pour lesquelles il n'existe pas de symboles normalisés spécifiques.

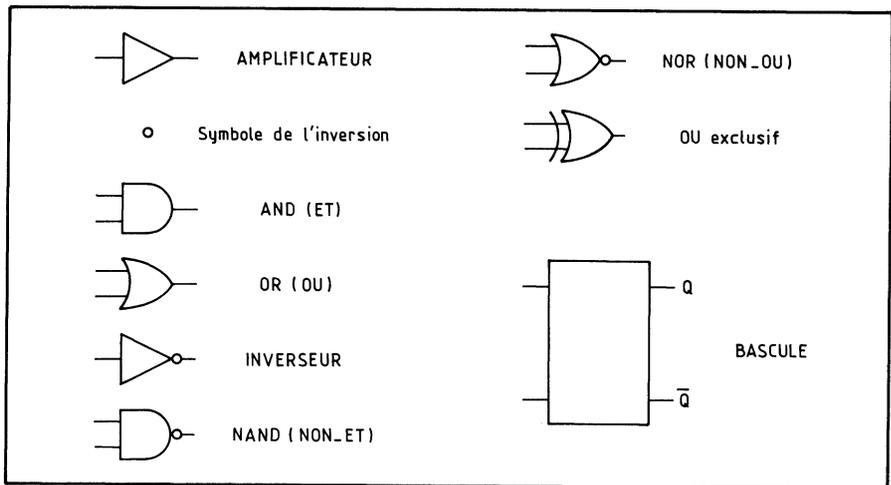


Fig. 7. — Les symboles les plus courants.

6. CARACTÉRISTIQUES PRINCIPALES DES CIRCUITS INTÉGRÉS

Les diverses technologies et familles mènent à des circuits intégrés dont les principaux paramètres caractéristiques sont :

- *Le temps de propagation du signal à travers une porte fondamentale*, noté t_{pd} : c'est le temps qui s'écoule entre le moment où un ordre est appliqué aux

entrées et l'instant où la décision est disponible à la sortie. Il se mesure généralement en nanosecondes (ns).

- *La consommation*, notée P_c (de puissance consommée) : c'est la puissance que prélève une porte fondamentale sur l'alimentation. Elle se mesure généralement en milliwatts (mW), en microwatts (μ W), et même parfois en nanowatts ($1 \text{ nW} = 10^{-9} \text{ W}$).

Cette puissance est dissipée sous forme de chaleur; c'est pourquoi on l'appelle également *puissance dissipée*. Elle intervient directement sur la densité d'intégration; en effet, chaque puce, ou chaque circuit intégré ne peut dissiper qu'une puissance limitée, ce qui oblige, lorsqu'on veut accroître la densité d'intégration, à recourir aux familles consommant le moins.

- *Le facteur de qualité*, F_Q , qui est le produit du temps de propagation par la consommation, et qui fournit un moyen de comparer les familles entre elles.

En toute rigueur et puisque l'on multiplie un temps par une puissance, on obtient un *travail* qui se mesure en *joules*, plus exactement en picojoules (pJ) s'il s'agit du produit de nanosecondes par des milliwatts.

Le tableau B compare ainsi quelques familles de circuits intégrés parmi les plus connues, à l'exception du *neurone*, qui est la cellule de base du cerveau.

- *La fréquence de travail* : elle dépend directement du temps de propagation, mais est généralement relevée sur des flip-flops, ce qui explique qu'il n'y ait pas de lien immédiat entre ces deux notions. Elle se mesure en mégahertz.

- *L'immunité au bruit* : elle traduit l'insensibilité du CI à tout ce qui n'est pas le signal utile, parasites transportés par les lignes, induits ou transitoires intempestives (de commutation, en particulier), chutes ohmiques dans les lignes (de masse, par exemple...).

Cette notion est difficile à appréhender car elle fait intervenir de nombreux paramètres : énergie du bruit, impédance aux bornes de laquelle il est appliqué, tension instantanée développée, durée du parasite, et même temps de montée et de chute caractérisant les circuits intégrés, et temps de propagation. On distingue ainsi :

- *l'immunité statique*, qui peut être évaluée en volts,
- et *l'immunité dynamique*.

A partir d'une fonction de transfert typique (tension de sortie en fonction de tension d'entrée), la norme CCTU 17-03 (fascicule 1 A) a défini une terminologie et dressé le gabarit de la figure 9 qui montre ce que l'on entend par immunité au bruit. On notera l'emploi des abréviations suivantes :

- H pour niveau *haut* (de *high*);
- L pour niveau *bas* (de *low*);
- I pour entrée (de *in*);
- O pour sortie (de *out*).

Par conséquent, on trouvera, si V est la tension :

- V_{IL} pour tension d'entrée au niveau *bas*;
- V_{IH} pour tension d'entrée au niveau *haut*;
- V_{OL} pour tension de sortie au niveau *bas*;
- V_{OH} pour tension de sortie au niveau *haut*.

- *La gamme des températures de service*, qui détermine les températures ambiantes qu'acceptent les circuits. On distingue généralement :

- la gamme « militaire », qui va de -55 à $+125$ °C;
- une gamme « commerciale », couvrant approximativement de -40 à $+85$ °C;
- la gamme « industrielle », plus restreinte puisqu'elle ne couvre que de 0 à $+75$ °C environ (selon les fabricants).

FONCTION	NORMES			EQUATION
	CIE	DIN 40 700	USA	
AND				$X = AB$
OR				$X = A+B$
NAND				$X = \overline{AB}$
NOR				$X = \overline{A+B}$
NAND avec 1 entrée inversée				$X = \overline{\overline{A}B}$
NOR avec 1 entrée inversée				$X = \overline{\overline{A}+B}$
Entrée d'inhibition				$X = (A+B) \overline{C}$
OU exclusif				$X = A\overline{B} + \overline{A}B$
Comparateur				$X = AB + \overline{A}\overline{B}$
ET câblé				
OU câblé				
Ligne à retard				
FLIP_FLOP				

Fig. 8. — Tableau des symboles, selon les normes citées.

- *La gamme des températures de stockage*, qui excède, elle, les valeurs ci-dessus; on aura, par exemple, de -65 à $+150$ °C.
- *Les valeurs des tensions d'alimentation* : bien des MOS exigent non pas seulement une, mais deux tensions d'alimentation. Par exemple, $+5$ et -12 V, référencés à la masse (0 V).
- *Les niveaux logiques*, c'est-à-dire l'éventail des tensions qui correspondent aux 1 et 0 logiques.
- *La sortance*, qui définit le nombre d'entrées que la sortie d'un circuit peut attaquer en parallèle, étant entendu qu'il s'agit toujours d'une seule et unique famille.

Temps de propagation et courbes caractéristiques

La figure 10 précise la définition du *temps de propagation*, t_{pd} , décomposé en ses temps fondamentaux. La courbe supérieure représente le signal d'entrée, la courbe inférieure, le signal de sortie supposé inversé. La tension d'alimentation est notée V_{DD} ; les notations sont d'origine américaine : r vient de *rise* (montée) et f de *fall* (chute), H vient de *high* (haut) et L de *low* (bas). Les définitions sont les suivantes :

- Temps de propagation, sortie *haut* vers *bas* : t_{PHL}
- Temps de propagation, sortie *bas* vers *haut* : t_{PLH}
- Temps de transition aux sorties : t_{THL} OU t_{TLH}
- Largeur d'impulsion (horloge), niveaux *bas* ou *haut* : t_{WL} OU t_{WH}
- Fréquence d'horloge : f_{CL} (*cl* pour *clock*, horloge)
- Temps de montée et de chute de l'horloge : t_{FCL} et t_{RCL}
- Temps d'établissement (*set-up*) : t_{SU}
- Temps de maintien (*hold time*) : t_H
- Temps d'exclusion (*removal time*) pour les signaux d'initialisation *set*, *preset* ou *reset* : t_{rem}

Dans le cas de circuits bipolaires, du type TTL par exemple, la référence de tension se situe généralement à mi-course de l'excursion globale de tension; une valeur type de 1,5 V a généralement été adoptée (la tension d'alimentation, elle, est de 5 V).

Dans le cas des circuits synchrones, l'horloge pourra être représentée comme le montre la figure 11. Enfin, l'action d'une commande de mise en 3^e état mène aux courbes de la figure 12.

7. LES FAMILLES DE CIRCUITS INTÉGRÉS

Les familles en logique bipolaire sont assurément les plus nombreuses. La figure 13 donne les structures de leurs portes fondamentales. Seules subsistent, en fait, la TTL dans ses variantes Schottky essentiellement, l'ECL, et la toute nouvelle venue qu'est l'I²L.

En MOS, les portes fondamentales sont développées figure 14. On remarquera, par comparaison, la simplicité des structures, même si ce ne sont pas les mêmes fonctions qui sont retenues.

La figure 15 montre l'évolution, extrapolée jusqu'en 1980, de la densité d'intégration (nombre de composants par circuits) et du prix par composant des circuits intégrés. On espère réussir à fabriquer, au début des années 1980, des circuits intégrés à un million de composants, au prix de 100 F actuels, soit environ 0,01 centime le composant!

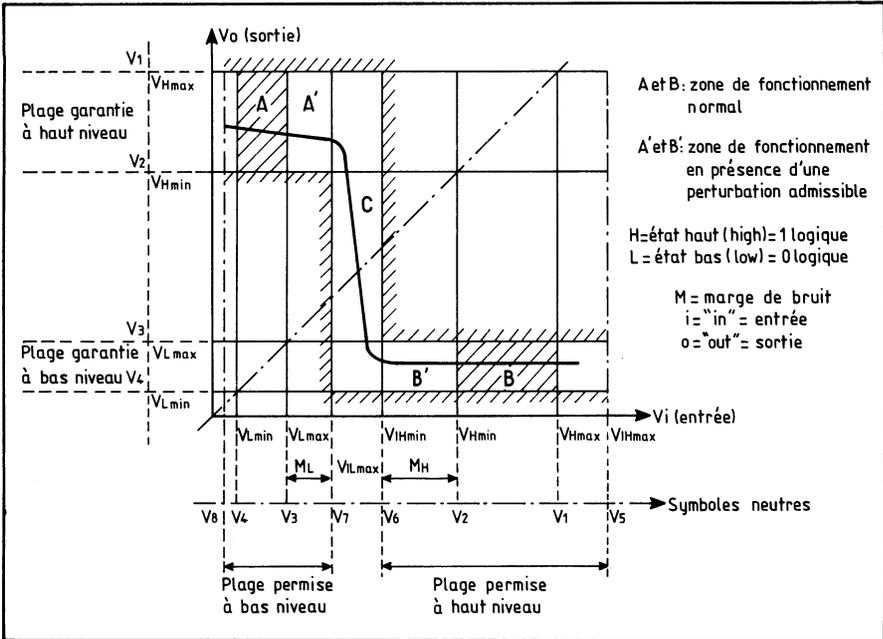


Fig. 9. — Gabarit des caractéristiques typiques de transfert, en tension, montrant les plages de fonctionnement normales et admises en présence de perturbations.

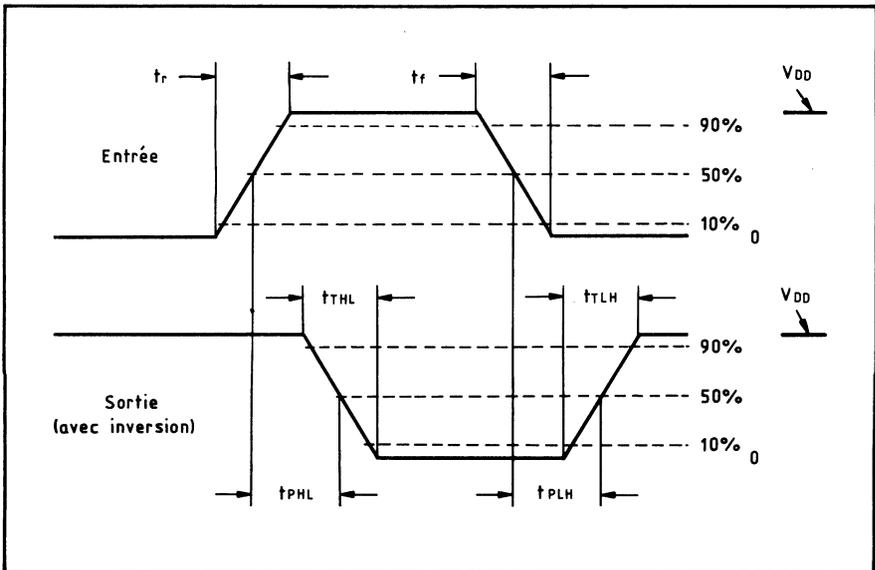


Fig. 10. — Courbes servant à définir les temps de propagation.

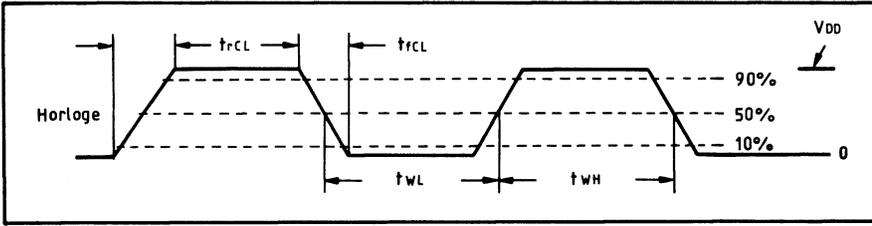


Fig. 11. — Définitions relatives aux signaux d'horloge.

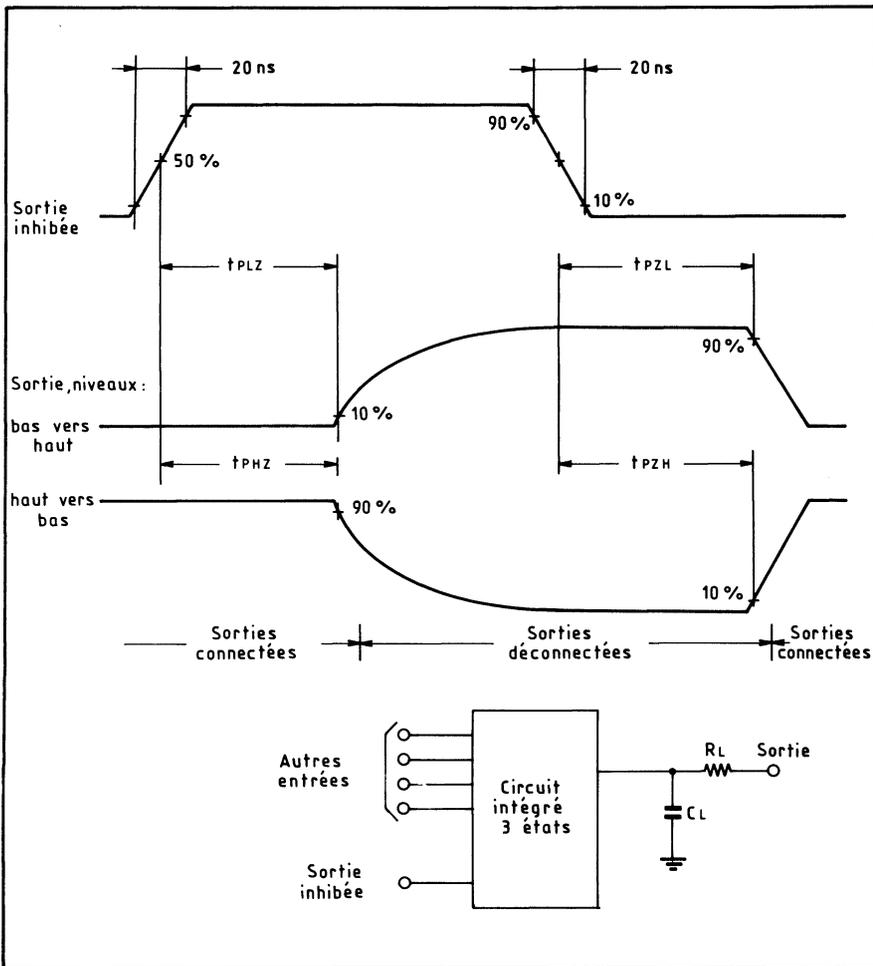


Fig. 12. — Le circuit à commande 3 états et ses temps.

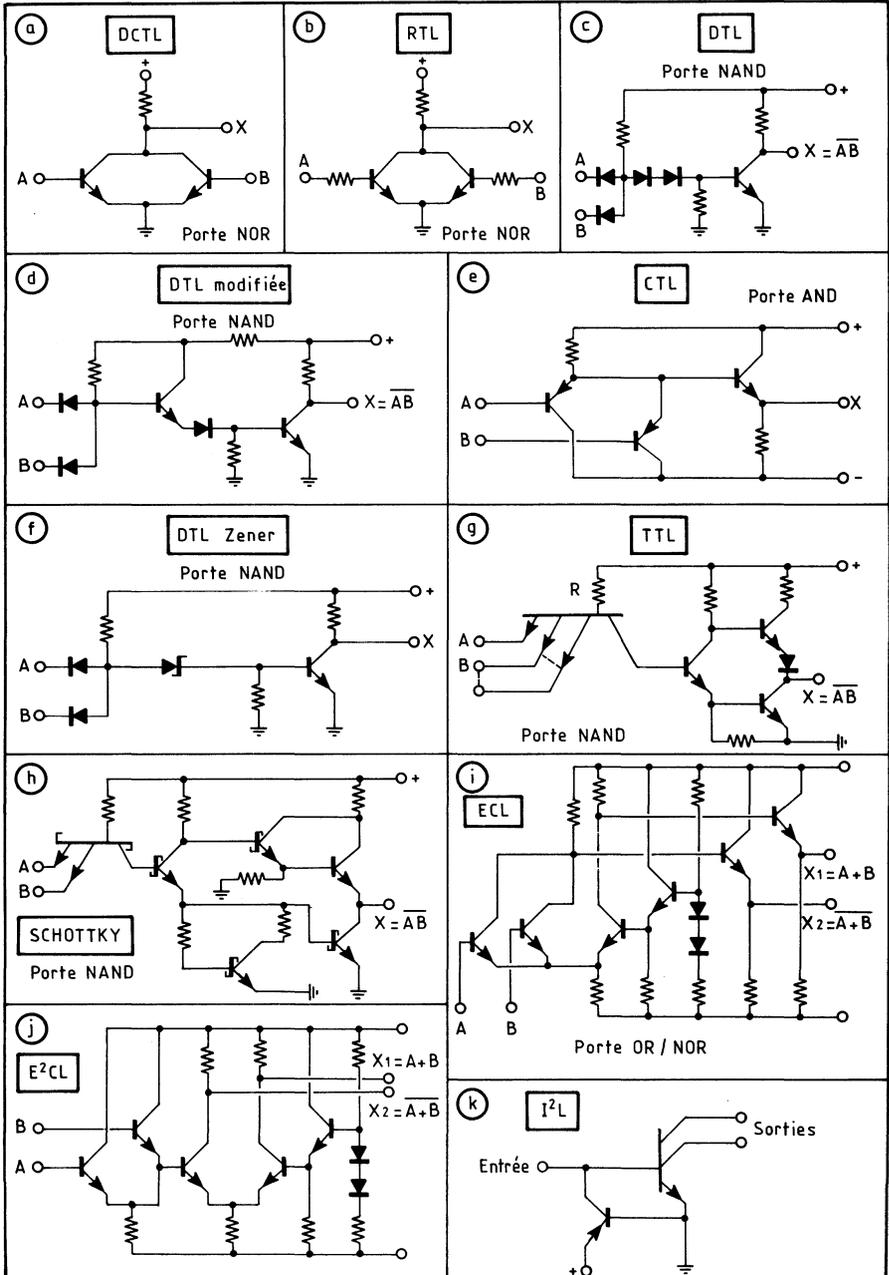


Fig. 13. — Structure des portes fondamentales dans les principales familles bipolaires.

**Tableau B. —
Quelques familles comparées :
caractéristiques de la porte fondamentale**

FAMILLE		t_{pd} (ns)	P_c (mW)	Q_F (pJ)	F (2) (MHz)
Bipolaires	RTL	20	6	120	5
	DTL	15	12	180	10
	TTL 54/74	10	10	100	30
	Schottky	3	20	60	120
	Schottky LS	8	2	16	40
	ECL 10 k	2	25	50	500
	ECL 100 k	1	5	5	800
	I²L	10	0,1	1	30
	CDI	5	2	10	40
MOS	PMOS	200	5	1000	5
	NMOS	60	5	300	10
	CMOS	60	1	60	10
	CMOS/SOS	3	1	3	
CTD	CCD (1)	50	$5 \cdot 10^{-3}$	0,25	10
	BBD (1)	100	$5 \cdot 10^{-3}$	0,5	5
	Neurone	$0,2 \cdot 10^6$	10^{-6}	0,2	

(1). Pour une cellule-mémoire et non une porte fondamentale.

(2).- Fréquence maximale pour un flip-flop.

8. L'HISTOIRE DES FET

L'histoire des FET commence, curieusement, bien avant celle des transistors bipolaires.

C'est en 1930, en effet, que le chercheur américain Julius Edgar LILIEFELD découvre qu'on peut commander la conduction dans un cristal en lui appliquant un champ électrique perpendiculaire. Il dépose alors plusieurs brevets relatifs à ce que l'on considère aujourd'hui comme le premier transistor à effet de champ, et dont le premier date, en réalité, de 1926 (*US Patent* n° 1900 018) et le second de 1930 (*US Patent*, n° 1745 175).

Puis un autre chercheur, britannique cette fois, s'intéresse à l'effet de champ : c'est Oscar HEIL, qui prend à son tour un brevet (*British Patent*, n° 439 457) en 1935; ses travaux ne sont suivis d'aucun développement pratique.

Il faut se reporter, ensuite, à la fin des années 30 : un jeune physicien des laboratoires *Bell*, William B. SHOCKLEY, cherche à remplacer les commutateurs électromécaniques utilisés en téléphonie pour établir les connexions; il s'intéresse alors aux dispositifs à l'état solide, et en particulier à une théorie proposée par Walter SCHOTTKY et relative à l'effet de redressement du courant alternatif constaté dans une jonction métal-semiconducteur.

SHOCKLEY entrevoit d'emblée la possibilité d'amplifier le signal par l'inter-

Fig. 14. — Trois inverseurs de base MOS.

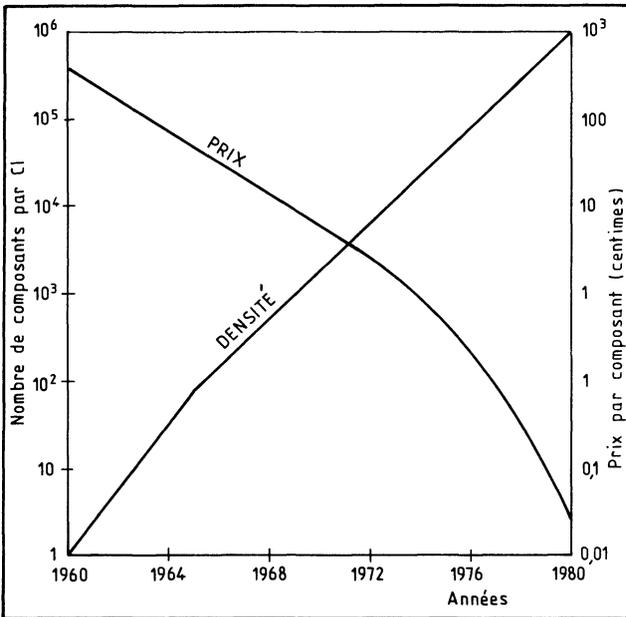
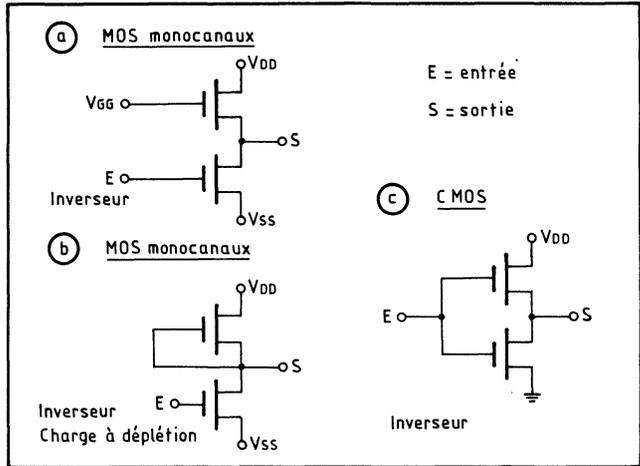


Fig. 15. — Évolution du prix par composant intégré et de la densité d'intégration, extrapolée jusqu'en 1980.

médiaire d'une zone à charge d'espace ou à déplétion (appauvrissement) qui s'étendrait dans le semiconducteur, sous le métal. Ses premiers essais pour réaliser un tel dispositif avec du cuivre et de l'oxyde de cuivre, en 1939, restent infructueux.

Il reprend cette étude après la seconde guerre mondiale avec John BARDEEN et Walter H. BRATTAIN en utilisant du germanium, semiconducteur mieux connu que la combinaison cuivre-oxyde de cuivre. Cette équipe travaille sur les potentiels de contact en surface et la charge d'espace en profondeur, et crée, en 1947, le premier transistor à pointes

L'année suivante en 1948, SHOCKLEY met au point le premier transistor bipolaire à jonctions. La même année, SHOCKLEY et G eral L. PEARSON, aux *Bell Laboratories*, observent l'action d'un champ sur une jonction PN de silicium. Puis, en 1952, SHOCKLEY publie la th orie du transistor   effet de champ tandis que George G. DACEY et Ian M. ROSS r alisent le premier  l ment, en 1953,   l'aide de germanium.

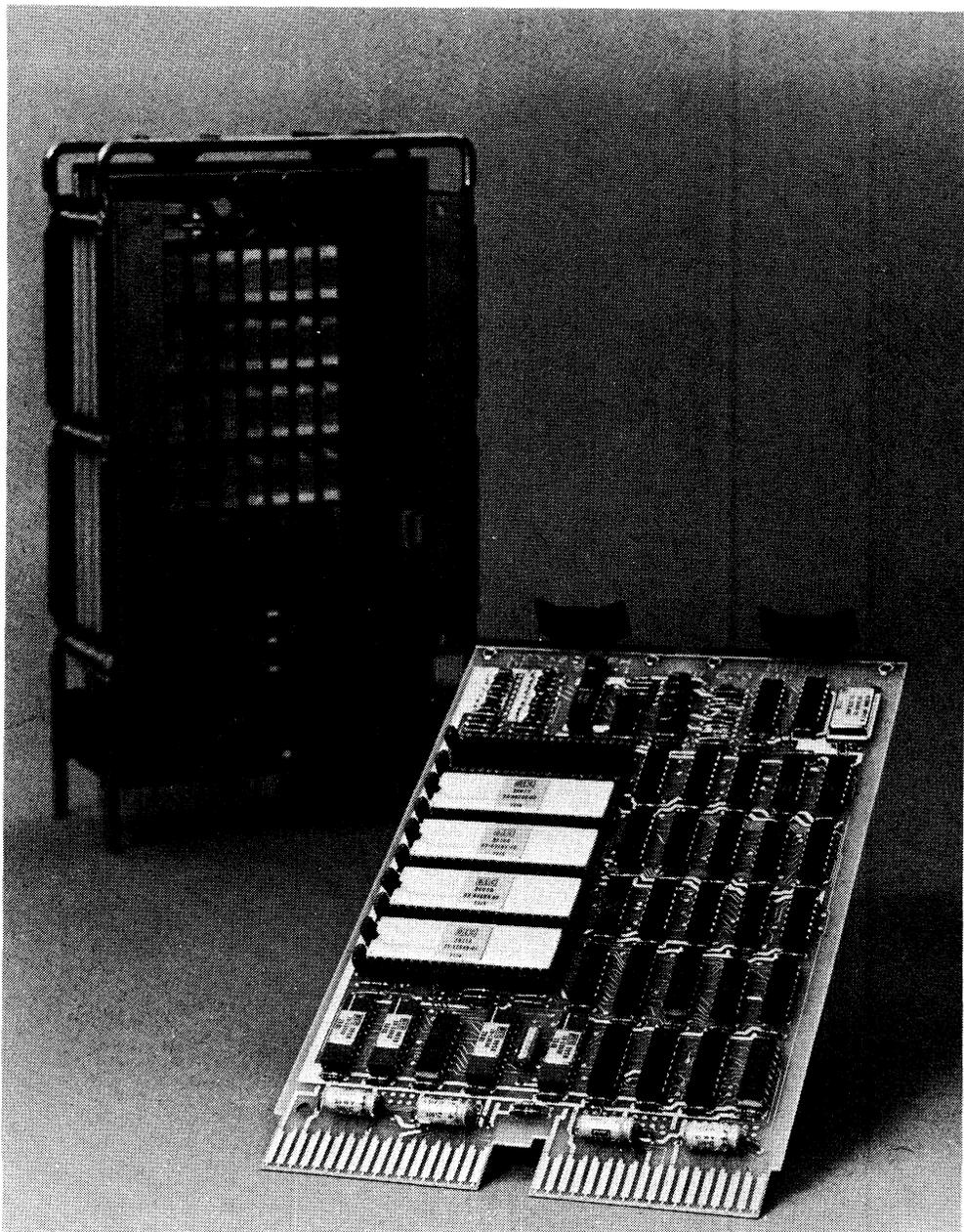
C'est  galement   l'aide de germanium qu'aux alentours de 1955 un chercheur fran ais du *Centre National d' tudes des T l communications* (CNET), Stanislas TESZNER, con oit le « Tecnetron », nom compos  des premi res lettres du nom de l'inventeur et des initiales du CNET. Le Tecnetron est un baton de germanium dans lequel a  t  pratiqu  un  trangement, une sorte de gorge tr s mince, sur lequel on r alise une jonction alli e polaris e en inverse, afin de r duire le courant traversant le barreau du germanium. C'est donc l  l'un des tout premiers transistors   effet de champ du monde, et qui fonctionne   des fr quences relativement importantes (30 MHz d s l'origine). Technetron sera, ainsi, synonyme de « Fieldistor », puisqu'ainsi ont  t  d sign s les premiers FET.

Mais le silicium commence   prendre le pas sur le germanium; sa gamme de temp rature est plus large et son traitement plus ais . En 1960, Dawon KAHNG et John ATALLA, des *Bell Laboratories*  galement, proposent une structure au silicium dans laquelle une  lectrode de commande isol e, ou gate, provoque la cr ation d'un canal conducteur entre des jonctions PN.

A peu pr s   la m me  poque, J. Torkel WALLMARK, alors   *RCA*, entrevoit la possibilit  d' laborer des fonctions logiques   l'aide de MOS. Un autre chercheur des laboratoires *RCA*, le D^r Paul K. WEIMER, d veloppe cette id e   l'aide de MOS   films minces de sulfure et de s l niure de cadmium.

Simultan ment, *Texas* d pose un brevet sur le circuit int gr  (1959), et *Fairchild* met au point le proc d  « Planar » (1960). D s lors, les premiers circuits int gr s apparaissent, bipolaires puis MOS. Le tout premier circuit int gr  MOS semble d    HOFSTEIN et HEIMAN, travaillant aux laboratoires de recherches d' lectronique,   *RCA*; il s'agit d'un r seau de huit paires de transistors   canal N, interconnect s pour former une double porte   quatre entr es.

Les travaux sur les FET porteront alors dans tous les azimuts : mat riaux nouveaux, tel l'AsGA (ars niure de gallium); structures nouvelles, telle le FET Schottky, etc.; S. TESZNER proposera encore le Gridistor et le VFET, puis na tront les VMOS (1976) et d sormais, les FET repousseront la barri re des puissances aux centaines de watts et celles des fr quences   la cinquantaine de gigahertz.



L'une des applications parmi les plus importantes des circuits MOS, l'informatique, illustrée ici par le micro-ordinateur LS1-11/2 de Digital Equipment. Le microprocesseur est constitué par les quatre circuits clairs de la carte; il traite des mots de 16 bits.

CHAPITRE II

FONCTIONNEMENT DU MOS FONDAMENTAL A ENRICHISSEMENT, CANAL P

1. PRINCIPE DE FONCTIONNEMENT

Pour présenter le transistor MOS, on se référera à un élément à enrichissement, à canal P. On verra plus loin ce que signifient ces caractéristiques.

Un MOS est constitué par un substrat de silicium N dans lequel sont diffusées deux régions P, appelées respectivement la source et le drain et notées S et D (fig. 1).

Au-dessus de l'espace drain-source est déposée une mince couche d'oxyde, de 0,1 à 0,2 μm d'épaisseur environ, métallisée en surface avec de l'aluminium pour constituer le gate, noté G. Les autres régions, non actives, sont recouvertes d'une couche épaisse d'oxyde. Le substrat est parfois désigné par la lettre B, initiale du mot américain « bulk » (corps, ou substrat). Polarisons ce MOS en appliquant (fig. 2) :

- la masse sur la connexion de source;
- une tension négative $-V_D$ sur le drain;
- une tension négative variable $-V_G$ sur le gate;
- la masse au substrat.

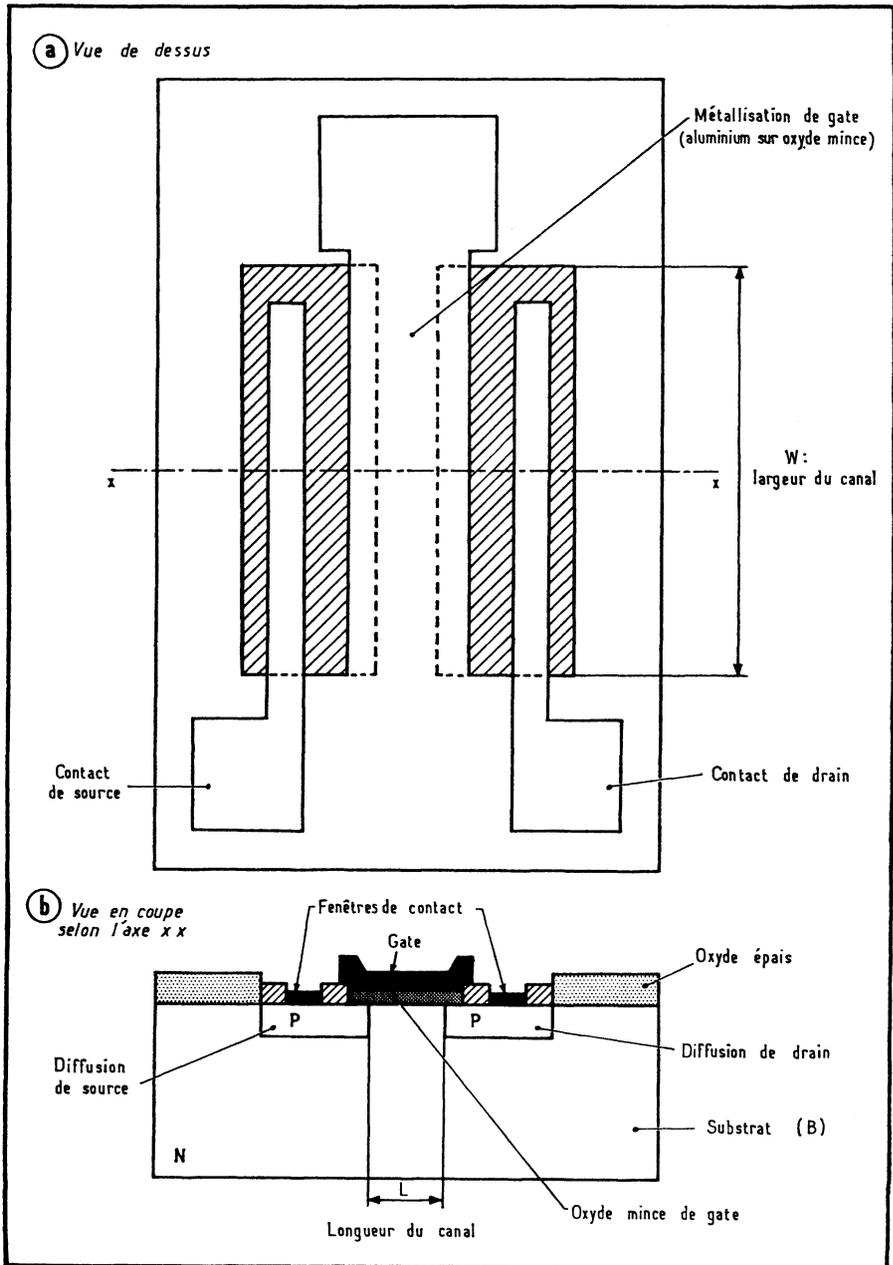


Fig. 1. — Vue de dessus (en a) et en coupe selon l'axe XX (en b) d'un MOS à canal P.

Faisons varier $-V_G$ à partir d'une tension nulle. Tant qu'un certain seuil négatif n'est pas atteint, aucun courant n'arrive au drain : I_D est nul et le MOS est dit *bloqué*.

La tension V_G évoluant négativement, elle atteint un niveau V_{TH} pour lequel le transistor entre en conduction. Ce niveau est la *tension de seuil*, le *threshold voltage* des Américains, ce qui explique l'abréviation utilisée : elle est de l'ordre de -4 V pour $I_D = 10 \mu A$, avec le dispositif fondamental examiné ici. Ajoutons que pour la représenter, on trouvera aussi bien les symboles V_{TH} que V_T tout simplement. V_G évoluant encore, le courant I_D prend l'allure de la courbe figure 2 b, et en résumé :

- pour $V_G > V_{TH}$, le MOS est bloqué;
- pour $V_G < V_{TH}$, le MOS est conducteur.

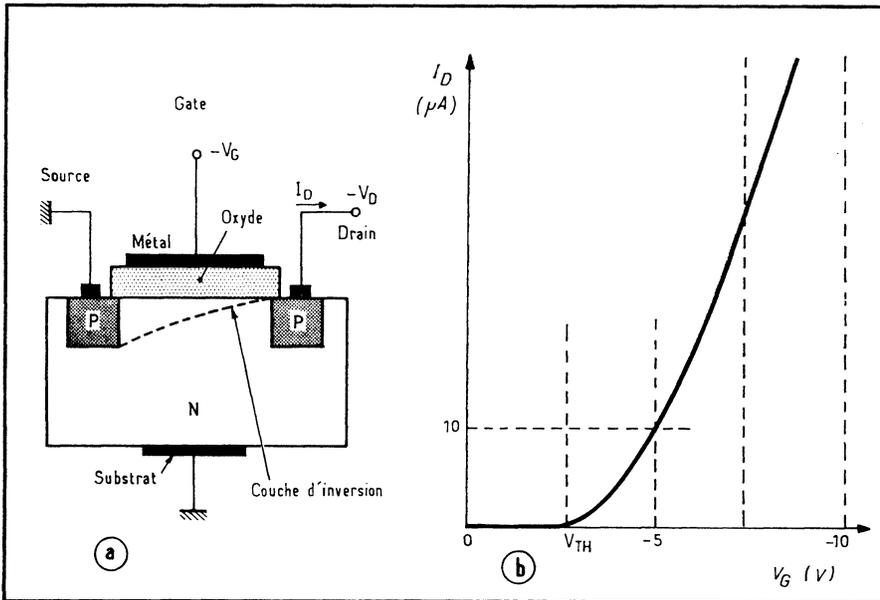


Fig. 2. — Conduction par enrichissement d'un MOS à canal P (a); caractéristiques I_D , V_D d'un MOS (b).

Ces valeurs sont évidemment des valeurs absolues, les tensions étant négatives rappelons-le.

Pourquoi une conduction se manifeste-t-elle? Parce que la zone située sous la gate, appelée *canal*, est *enrichie* en charges positives sous l'action de la polarisation croissante de la gate. En effet, si une polarisation négative est appliquée au gate, elle chasse les électrons libres du substrat sous-jacent mais attire, par contre, les charges positives, ou *trous*, P.

Quand suffisamment de trous se sont accumulés en surface, cette zone qui était initialement N tend à devenir P : il y a donc inversion des porteurs. Et si $-V_G$ est suffisante, il se forme un *canal P* entre le drain et la source qui assure la conduction dans le MOS (fig. 3). Le courant qui traverse le transistor peut désormais emprunter un chemin continu P : source, canal et drain.

On notera que la polarité du canal est (toujours) l'inverse de celle du substrat.

Action de la tension de drain. Zones triode et saturée

Grâce à la tension V_G , on parvient ainsi à commander la conduction. Mais on peut également faire varier la forme du canal en agissant sur la tension V_D (ou plus exactement sur la différence de tension drain-source, notée V_{DS}). Ainsi, en maintenant constante V_G et en variant V_D , on aboutit aux différentes formes de canal de la figure 4 :

- si V_D est nulle et si V_G est suffisamment négative, la couche d'inversion le long du canal P est celle représentée en *a*; elle est pratiquement uniforme;
- V_D devenant négative, le canal commence à « s'étrangler » (en *b*). Il se rapproche de la surface près du drain car la différence de tension gate-substrat détermine la largeur de la couche d'inversion; or, cette différence passe de $V_G - 0$ près de la source à $V_G - V_D$ près du drain. C'est la *région triode*. Elle est ainsi appelée car le dispositif se comporte alors approximativement comme une résistance ohmique, et sa courbe caractéristique (fig. 5) est fort semblable à celle des anciennes lampes triodes;

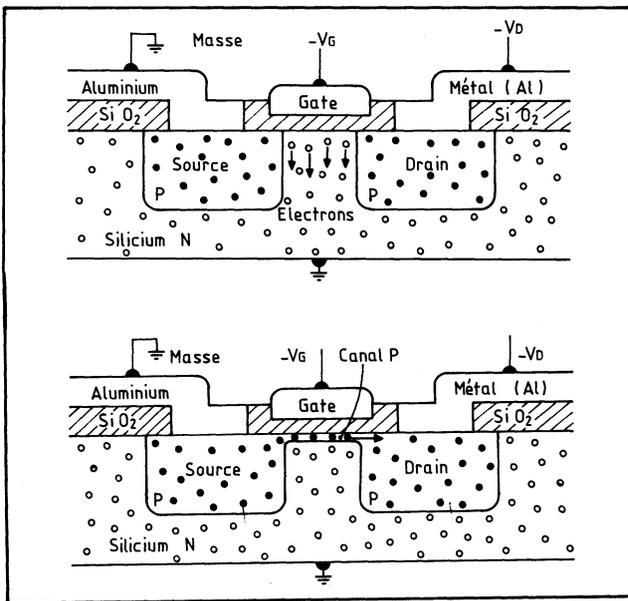


Fig. 3. — Distribution des porteurs et création du canal. Les trous sont les points pleins, les électrons des cercles.

- V_D devenant encore plus négative, la couche d'inversion disparaît près du drain (en *c*). Le moment précis où le canal cesse de relier la source au drain est obtenu lorsque $V_G - V_D = V_{TH}$, tension de seuil;
- puis, V_D s'accroissant encore négativement, le canal est franchement interrompu (en *d*). A ce moment on a : $V_G - V_D < V_{TH}$. Cependant, et c'est là un phénomène important, un courant pratiquement indépendant de la tension source-drain peut encore circuler entre ces deux électrodes, du fait de la déplétion (de l'appauvrissement) qui existe dans la jonction inverse drain et substrat (ce dernier étant, rappelons-le, connecté à la masse).

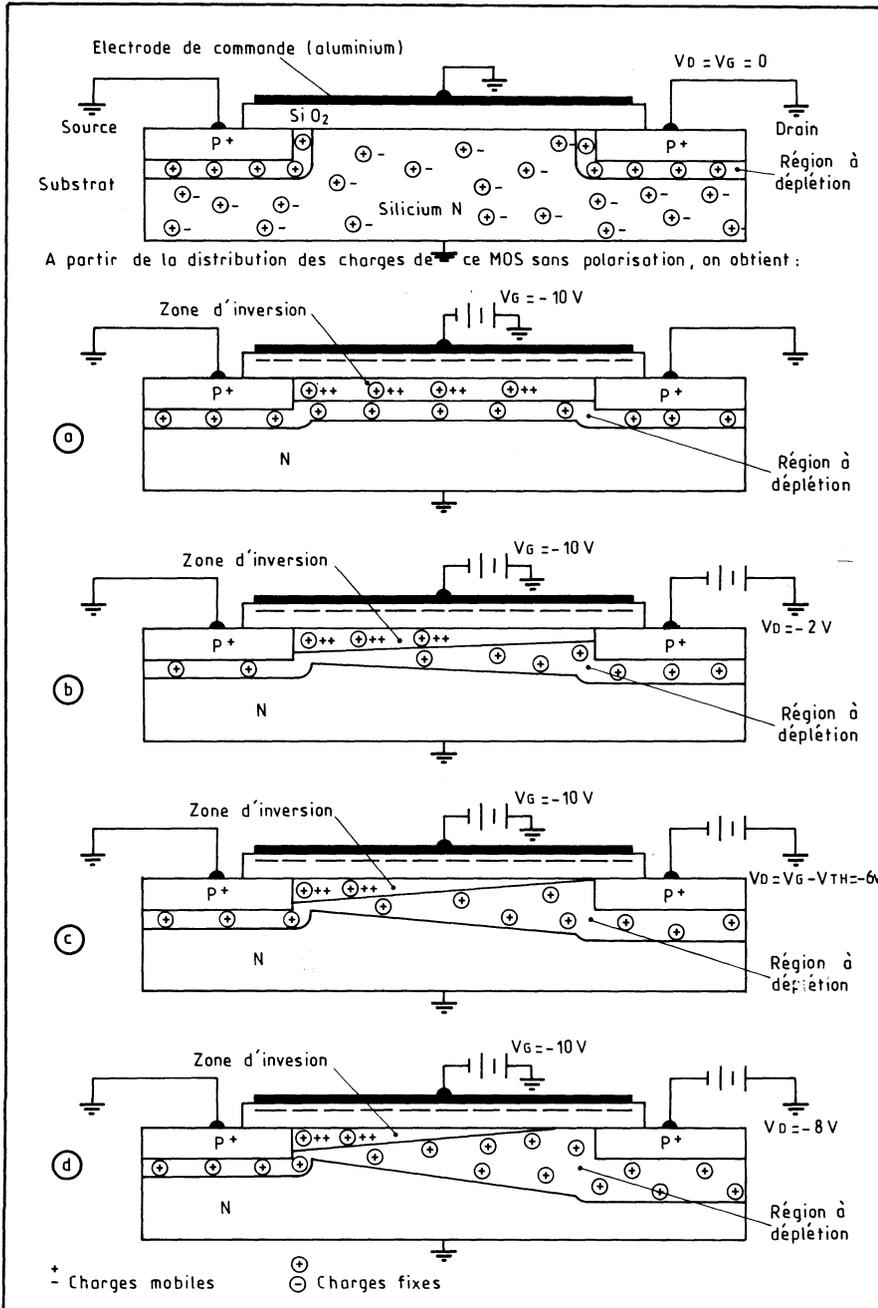


Fig. 4. — Quelques formes du canal dans un MOS pour différentes valeurs de la tension de drain.

Examinons en effet ce qui se passe le long du canal : le potentiel n'est pas constant et varie, de la source au drain, en fonction de la tension de drain. Il est donc évident que la tension de commande V_G agit différemment le long du canal puisque la différence de tension $V_G - V_{\text{canal}}$ n'est pas constante; il s'ensuit que la densité des porteurs décroît le long du canal et est minimale à proximité du drain.

Quand le courant drain s'est accru au point qu'il n'existe plus de porteurs mobiles près du drain, il y a *saturation*. Si l'on accroît la densité des porteurs en élevant V_{GS} , on augmente la valeur du courant de saturation; par contre la tension V_D n'agit pas sur la densité des porteurs mais sur la fraction de longueur du canal soumise à saturation; elle est donc pratiquement sans effet sur le courant de saturation.

Pour illustrer le phénomène de saturation, on dit que le canal « se pince »; sa longueur diminue. Toute la tension V_{DS} apparaît aux bornes de la zone de canal non pincée et n'agit plus sur la valeur du courant.

On dit qu'on travaille dans la zone *saturée*, ou *zone pentode* par analogie avec la courbe typique des anciens tubes pentodes. Le dispositif se comporte alors comme une source de courant I_D indépendant de V_{DS} , cette dernière pouvant être relativement élevée.

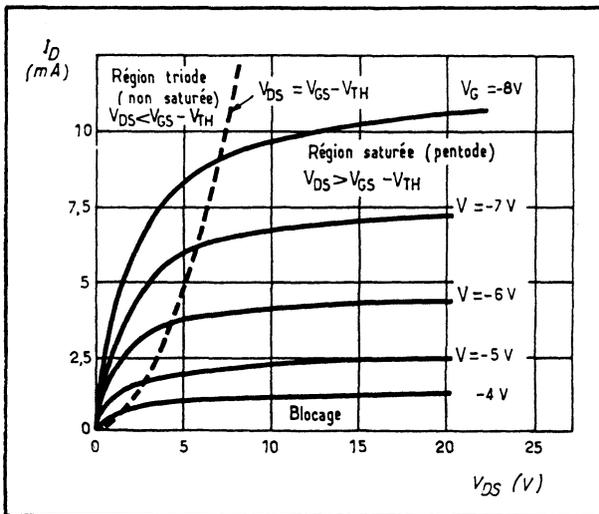


Fig. 5. — Caractéristiques V_{DD} , I_{DD} en fonction de la tension V_{GS} .

Le réseau classique des caractéristiques du MOS donnant le courant de drain I_D en fonction de la tension drain-source V_{DS} pour différentes valeurs de la tension gate-source V_{GS} a ainsi été tracé figure 5.

En conclusion, les conditions sont, en référant les tensions à la source :

- | $V_{GS} - V_{TH} | = | V_{DS} |$ zone de transition
- | $V_{GS} - V_{TH} | \leq | V_{DS} |$ région saturée (pentode)
- | $V_{GS} - V_{TH} | \geq | V_{DS} |$ région triode (non saturée)

On notera encore un fait important : le drain, polarisé négativement, constitue avec le substrat une diode polarisée en inverse (donc bloquée). D'autre part, et avec la création du canal, il se forme dans le substrat une couche d'équilibre, intermédiaire entre le canal P et le substrat N, d'où les porteurs sont quasiment

exclus, qu'ils soient P ou N. Cette couche est appelée *couche d'appauvrissement* (ou de *déplétion*), et représentée en blanc sur la figure 6. De ce fait, le transistor MOS élémentaire se trouve donc bien isolé électriquement du substrat, et donc d'un éventuel MOS voisin : il y a là une différence considérable avec les bipolaires qui ne sont pas, eux, isolés « naturellement » et qui exigent que l'on crée des « caissons d'isolement » spéciaux source d'opérations supplémentaires et qui occupent une surface non négligeable de silicium.

2. THÉORIE DU FONCTIONNEMENT DU MOS

La capacité de gate

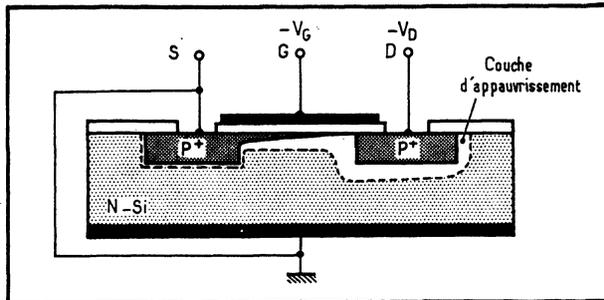
L'électrode métallique du gate forme, avec le substrat, un condensateur dont le diélectrique est la silice (dioxyde de silicium : SiO_2). Sa capacité C_{Ox} est donnée de façon classique par :

$$C_{\text{Ox}} = \frac{\epsilon_0 \epsilon_r S}{d}$$

où :

- ϵ_0 est la permittivité du vide et vaut $8,86 \cdot 10^{-12} \text{Fm}^{-1}$;
- ϵ_r est la permittivité relative de la silice et vaut environ 3,9.

Fig. 6. — En dessous des électrodes du MOS se forme une zone dépourvue de porteurs de charge et appelée couche d'appauvrissement (ou de déplétion) ou encore couche désertée.



(On trouvera, dans les tableaux A et B, quelques caractéristiques utiles relatives au silicium, à la silice, ainsi qu'au germanium et à l'arséniure de gallium.)

• S est la surface, égale au produit WL de la figure 2-1; la longueur L est généralement comprise entre 6 et 10 μm pour le MOS fondamental, alors que la largeur W peut varier entre 20 et 2000 μm . Prenons ici, pour les besoins de l'exemple, une valeur importante de S, soit 8 000 μm^2 .

• d (qu'on note aussi souvent t_{Ox}) est l'épaisseur de l'oxyde, de l'ordre de 150 nm.

Cette capacité est ainsi de l'ordre de :

$$C_{\text{Ox}} = \frac{8,86 \cdot 10^{-12} \cdot 3,9 \cdot 8 \cdot 10^3}{1,5 \cdot 10^{-7}} \approx 1,8 \text{ pF}$$

Pour étudier de plus près le comportement du MOS, on va se servir non pas de la capacité réelle C_{Ox} mais de la capacité par unité de surface, qui sera notée C_o . On a alors :

$$C_o = \frac{\epsilon_o \epsilon_r}{d}$$

Soit :

$$\begin{aligned} &= 2,3 \cdot 10^{-4} \text{ F/m}^2 \\ &= 230 \text{ pF/mm}^2 \end{aligned}$$

Tableau A. — Principales caractéristiques à 300 °K

CARACTÉRISTIQUES	Si	Si O ₂	Ge	As Ga
Bande interdite (eV)	1,11	8	0,67	1,4
Densité effective d'états énergétiques :				
• Bande de conduction (at.cm ⁻³)	2,8.10 ¹⁹		1,04.10 ¹⁹	4,7. 10 ¹⁷
• Bande de valence (at.cm ⁻³)	1,04.10 ¹⁹		0,6.10 ¹⁹	7.10 ¹⁸
Concentration intrinsèque de porteurs par cm³	1,45.10 ¹⁰		2,4.10 ¹³	9.10 ⁶
Mobilité pour matériau intrinsèque :				
• Electrons (cm ² /Vs)	1350	isolant	3900	8600
• Trous (cm ² /Vs)	480		1900	250
Constante diélectrique	11,7	3,9	16,3	12
Point de fusion (°C)	1415	~ 1700	937	1238
Champ de claquage (V/μm) ≈	30	600	8	30

La tension de seuil

Pour développer la théorie du MOS, on fait intervenir les données suivantes :

- *Le nombre de porteurs par unité de surface du canal*, et ses relations avec les tensions V_{DS} , V_{GS} et le potentiel du canal. On considère en effet que la conduction s'effectue pratiquement en surface du canal, et non en profondeur, ce qui paraît évident *a priori* compte tenu du principe de commande du canal.
- *La mobilité de ces porteurs dans le champ drain-source.*

On est amené à définir, en outre :

- Q_{SS} , la densité de charges par unité de surface, piégée dans l'oxyde à l'interface oxyde-silicium. C'est une valeur fixe;

Tableau B. — Quelques grandeurs pratiques ou typiques relatives au MOS « classique » P à enrichissement, à 300 °K

GRANDEUR	SYMBOLE	VALEUR	UNITÉ
Mobilité des électrons dans le silicium	μ_n	0,15	$m^2/(Vs)$
Concentration intrinsèque en porteurs dans le Si	n_i	$1,4 \cdot 10^{16}$	m^{-3}
Charge de l'électron	q	$1,6 \cdot 10^{-19}$	C
Capacité de gate par unité de surface pour oxyde de 100 nm	C_0	350	$\mu F/m^2$
Mobilité des trous dans le canal			
• silicium 1.1.1	μ_p	$17 \cdot 10^{-3}$	$m^2/(Vs)$
• silicium 1.0.0	μ_p	$13 \cdot 10^{-3}$	$m^2/(Vs)$
Potentiel de contact métal-silicium	ϕ_{MS}		
• Si P		0,9	V
• Si N		0,1 à 0,4	V
Résistivité intrinsèque du silicium	ρ	63 000	$\Omega \cdot cm$

- Q_B , la charge fixe dans la zone de déplétion située sous la zone d'inversion et due aux atomes ionisés résultant du dopage du substrat;
- Q_P , la charge mobile dans la couche d'inversion, donc des trous (d'où l'indice P) dans notre cas, ou :
- Q_N s'il s'agit d'électrons;
- enfin Q_G , charge dans l'électrode de gate qui équilibre la somme des charges ci-dessus.

Approximation de 1^{er} ordre

D'autre part, et pour développer la théorie du MOS, on recourt à plusieurs simplifications. La théorie la plus simple, qui ne tient pas compte de la charge dans la zone de déplétion du MOS, est ainsi appelée théorie de 1^{re} approximation ou 1^{er} ordre d'approximation; elle suppose que :

1. La charge induite dans le canal est due uniquement à la tension de gate.
2. La charge Q_{SS} présente est constante.
3. Le profil de dopage est uniforme dans l'épaisseur du substrat.
4. La conduction survient entre drain et source lorsque la densité des porteurs mobiles en surface égale celle des porteurs majoritaires dans le substrat.
5. La conduction s'effectue par courant de diffusion.
6. Une mobilité moyenne des porteurs, en surface, peut être définie indépendamment des tensions de drain et gate.

La tension de seuil est alors définie comme étant la tension gate-source V_{GS} à partir de laquelle la conduction s'établit et croît ensuite très vite. On suppose encore que le substrat et la source sont au même potentiel (fig. 7), et que la tension drain-source est faible. La distribution des différentes charges définies ci-dessus peut être illustrée par la figure 8, où l'on a symbolisé les charges dans le cas où il y a accumulation d'électrons (porteurs N), pas de porteurs (bande plate), une zone de déplétion, et enfin une zone d'inversion due aux trous P.

La tension de seuil V_{TH} dépend alors, selon la théorie de première approximation :

1. Du niveau de dopage N_D du substrat, dont dépend le potentiel de Fermi Φ_F . Celui-ci est de l'ordre de $-0,30$ à $-0,25$ V; on trouvera par exemple, $0,25$ V pour du silicium dopé N de $5 \Omega\text{cm}$ de résistivité, et pour $N_D = 10^{15} \text{ cm}^{-3}$.

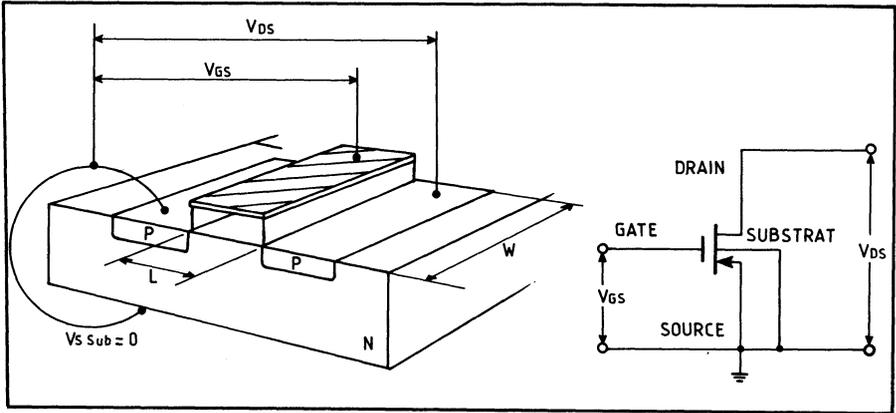


Fig. 7. — Tensions appliquées au MOS lorsque le substrat est à la masse.

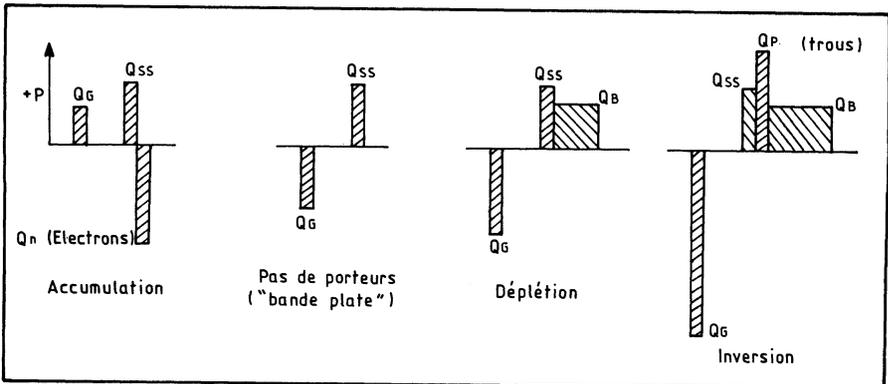


Fig. 8. — Distribution possible des charges.

2. De la charge fixe dans la zone de déplétion, Q_B ; en surface, on peut considérer qu'on trouve $2 \cdot 10^{11}$ charges électroniques élémentaires pour le même niveau de dopage, par centimètre carré.

3. De la charge fixe en surface, dans l'oxyde, soit Q_{SS} . On trouve quelque $1,5 \cdot 10^{11}$ charges électroniques par centimètre carré.

4. Du potentiel de contact Φ_{MS} qui exprime la différence de travail à fournir pour extraire un électron du conducteur de gate et un électron du substrat. Si le gate est en aluminium (notre cas), on trouve que :

$$\begin{aligned}\Phi_M - \Phi_s &= \Phi_{MS} = 3,2 - (3,8 + \Phi_F) \\ &= 3,2 - (3,8 - 0,25) = -0,35 \text{ V}\end{aligned}$$

5. Enfin, de C_O défini antérieurement, de l'ordre de $2,3 \cdot 10^{-8}$ F/cm².

La formule donnant V_{TH} est alors :

$$V_{TH} = -\frac{Q_{SS}}{C_O} - \frac{Q_B}{C_O} + \Phi_{MS} + 2\Phi_F$$

$$= -1,05 - 1,4 - 0,35 - 0,5 = -3,3 \text{ V}$$

Cette formule servira ultérieurement à montrer les répercussions des diverses technologies sur la tension de seuil.

Le courant de drain

Région triode

L'équation donnant le courant de drain I_{DS} (I drain-source) en fonction des tensions est :

$$I_{DS} = -\mu_p C_O \frac{W}{L} \left[(V_{GS} - V_{TH}) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

où μ_p est la mobilité des trous dans le canal.

Le produit $\mu_p C_O$ est souvent désigné par K (plus précisément, par K'). Ainsi, on a :

$$I_{DS} = -K \frac{W}{L} \left[(V_{GS} - V_{TH}) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Région pentode

Dans la région pentode (saturée), la valeur du courant de drain ne dépend plus de la tension V_{DS} . Elle est donné par :

$$I_{DS} = -\frac{K W}{2 L} (V_{GS} - V_{TH})^2$$

Il s'agit, en effet, d'un courant de saturation qu'on peut noter $I_{DS} (sat)$. Or, puisque la condition de saturation est : $V_{DS} \geq V_{GS} - V_{TH}$, on peut encore écrire dans le cas où $V_{DS} = V_{GS} - V_{TH}$

$$I_{DS} = -\frac{K W}{2 L} (V_{DS})^2$$

Cette tension $V_{DS} (sat)$ minimale détermine ainsi le passage à la saturation avec un courant $I_{DS} (sat)$.

Exemple

Soit à calculer le courant de saturation (région pentode) d'un MOS, avec :

- $W/L = 15$
 - K , qui est typique et constant pour un procédé technologique donné, varie généralement de 1 à $8 \cdot 10^{-6}$ A/V². Soit, ici, $K = 3 \cdot 10^{-6}$.
 - $V_{TH} = 4$ V
 - $V_{GS} = 12$ V
- On trouve :
- $$I_{DS} = -\frac{1}{2} \cdot 3 \cdot 10^{-6} \cdot 15 (12 - 4)^2$$
- $$\simeq -1,5 \text{ mA} \quad (\text{région saturée})$$

Si l'on veut calculer le courant dans la région triode, il faut faire intervenir V_{DS} , toujours faible. Soit $V_{DS} = 2$ V; on trouve

$$I_{DS} = (\text{région triode}) = -3 \cdot 10^{-6} \cdot 15 \left[(12 - 4) 2 - \frac{2^2}{2} \right]$$

$$\simeq 0,6 \text{ mA}$$

Remarques sur les coefficients K

Le choix des symboles n'ayant pas toujours été très heureux, on trouvera dans la littérature de base des coefficients K' et même K différents, par définition, et tels que :

et

$$K' = \frac{\mu_p \epsilon_0 \epsilon_r}{2d}$$

$$K = K' \frac{W}{L}$$

En conséquence, ce K' (et ce K) étant déjà divisés par 2, les formules ci-dessus doivent être multipliées par 2.

De la même façon, les notations pourront varier et l'on trouvera ϵ_{OX} pour ϵ_r , ainsi que t_{OX} pour d , épaisseur de l'oxyde; parfois encore, ϵ_{OX} sera le produit $\epsilon_0 \epsilon_r$, ce qui mène à $C_{OX} = \epsilon_{OX}/t_{OX}$, etc.

On se méfiera encore de la confusion suivante : *par définition*, V_{TH} est la tension à partir de laquelle le canal commence à se former. Par convention, le V_{TH} pratique est généralement mesuré à 1 μ A de courant drain et dépend alors de la géométrie du transistor, qui n'intervenait pas dans le cas précédent. Donc, ces valeurs pourront être sensiblement différentes.

Le rapport W/L et le facteur de gain β

On notera que le courant dépend du rapport W/L , largeur sur longueur du canal. Ce rapport joue donc un rôle important dans l'élaboration du MOS, et ce d'autant plus que L , longueur du canal entre source et drain, est pratiquement imposée par la technologie adoptée et varie ainsi de 10 à 2 μ m. Comme ce sont là des valeurs minimales résultant de la définition propre aux circuits, la largeur W ne pourra être qu'égale ou supérieure.

Le coefficient K , lui, est typique d'une technologie. Enfin, le produit KW/L est souvent appelé *facteur de gain*, et est alors désigné par la lettre grecque β .

Stabilité en température

La température affecte la tension de seuil, qui diminue avec son accroissement (fig. 9). Il en va d'ailleurs de même avec le coefficient K (fig. 10), lequel s'exprime en ampères par volt carré ou, ce qui revient au même, en mhos par volt.

La mobilité des porteurs P

Dans l'élaboration des formules précédentes intervenait la mobilité μ des porteurs positifs dans la zone d'inversion (le canal), qu'on désigne par μ_p ou μ_o . Les mesures expérimentales montrent qu'elle est de l'ordre de :

Fig. 9. — La tension de seuil en fonction de la température.

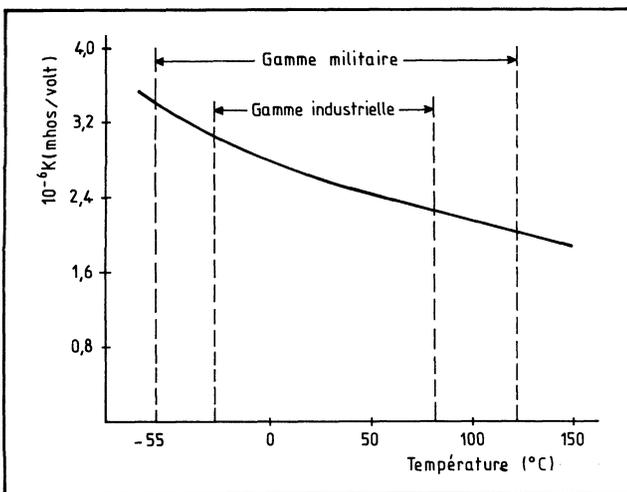
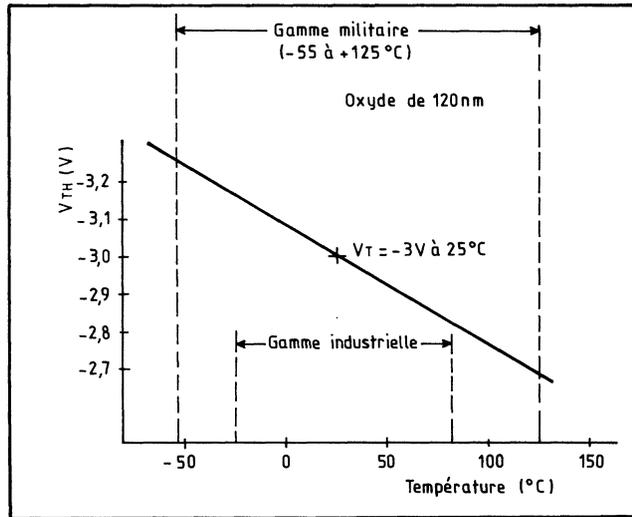


Fig. 10. — Le coefficient K en fonction de la température.

- $\mu_o = 170 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ dans le silicium orienté 1.1.1;
- $\mu_o = 130 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ dans le silicium orienté 1.0.0.

On verra, dans le chapitre consacré aux technologies, ce que traduisent ces plans 1.1.1 et 1.0.0.

Notons encore que, dans les mêmes conditions, la mobilité de porteurs négatifs (N) que sont les électrons serait de l'ordre de 500 à 600 $\text{cm}^2/(\text{Vs})$.

La résistance drain-source

La résistance drain-source du MOS, qui n'est autre que la résistance offerte par le canal, dépend essentiellement, pour une technologie déterminée :

1. De la géométrie du canal, donc du rapport W/L ;
2. Des tensions appliquées.

En toute rigueur, et pour V_{GS} constant, on a :

$$R_{DS} = \frac{dV_{DS}}{dI_{DS}}$$

ce qui donne dans la zone triode :

$$R_{DS} = \frac{1}{K \frac{W}{L}} \cdot \frac{1}{\left(V_{GS} - V_{TH} - \frac{V_{DS}}{2} \right)}$$

soit, puisqu'ici V_{DS} est faible, en simplifiant :

$$R_{DS} = \frac{1}{K \frac{W}{L}} \cdot \frac{1}{(V_{GS} - V_{TH})}$$

On constate que cette résistance est indépendante (pratiquement) de la tension V_{DS} ; de ce fait, on peut tirer une conclusion immédiate : le MOS peut servir de résistance, sa valeur étant commandée par la tension de gate V_{GS} . On verra, en effet, comment le MOS est utilisé en résistance de charge.

Exemples de valeurs

La résistance d'un MOS à l'état passant peut varier dans de très grandes limites, en fonction de sa géométrie et des autres caractéristiques. Ainsi, R_{DS} peut atteindre jusqu'à plusieurs kilohms, les valeurs les plus basses se situant autour de la centaine d'ohms généralement; cette résistance à l'état conducteur est donc supérieure à celle des transistors bipolaires.

Par exemple, à 25 °C, un MOS de rapport $W/L = 100$, à substrat de résistivité $3 \Omega \cdot \text{cm}$ avec un $V_{TH} = -3 \text{ V}$ aura une $R_{DS} (on)$ (R_{DS} à l'état passant) de l'ordre de 150 Ω .

Il est difficile de réduire cette valeur, sauf si l'on accroît le rapport W/L . Ainsi, pour obtenir une R_{DSon} de 15 Ω seulement, avec $V_{GS} = 10 \text{ V}$, il faudrait un rapport W/L de 800, soit pour un canal de 8 μm , une largeur de 6 400 μm . La surface occupée par un tel MOS serait de 0,2 mm^2 . Par contre, un rapport W/L de 5 mènerait à un R_{DS} de plus de 10 $\text{k}\Omega$.

Enfin, la résistance du canal varie avec la tension gate-source selon la région de fonctionnement, comme le montre la courbe de la figure 11 relevée sur un MOS type MEM 2009 de *General Instrument*, avec un courant drain de 0,1 mA. On notera que cette résistance intervient, en particulier, dans la constante de temps RC du dispositif (donc agit sur la bande passante), ainsi que lorsque les courants sont variables (fonctionnement analogique).

La transconductance

La transconductance, notée g_m , ou pente, est donnée à V_{DS} constant par :

$$g_m = \frac{dI_{DS}}{dV_{GS}}$$

Dans la zone saturée, et en reprenant la formule de I_{DS} , on trouve :

$$g_m (\text{saturée}) = K \frac{W}{L} (V_{GS} - V_{TH})$$

C'est donc l'inverse de R_{DSon} dans la zone triode.

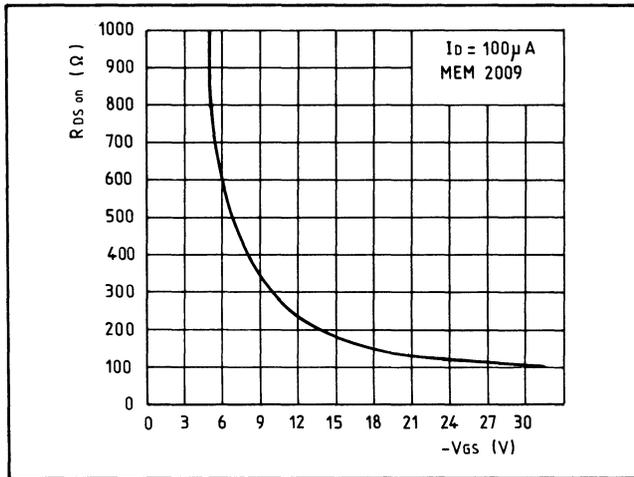


Fig. 11. — La résistance du canal varie avec la tension gate-substrat.

Dans la zone triode et au prix de simplifications maximales, on a approximativement :

$$g_m (\text{triode}) \simeq K \frac{W}{L} V_{DS}$$

Les valeurs typiques de pente s'étendent de 100 μmhos à 3 000 μmhos .

Le tableau de la figure 12 résume ainsi les caractéristiques du MOS fondamental selon la théorie de premier ordre.

Théorie du 2^e ordre

Avec la théorie du *second ordre*, on tient compte du fait que la charge de la zone de déplétion, située sous le canal, dépend du potentiel du canal. On supposait antérieurement qu'elle était constante (théorie de première approximation).

On a supposé, en effet, que le substrat et la source étaient reliés au même potentiel (0 V). Or, s'il est vrai que le substrat est systématiquement connecté au potentiel le plus positif, pour interdire toute polarisation en direct des diodes substrat-source, cela n'est plus exact des sources. En effet, si deux MOS sont connectés en série, par exemple, la source de l'un sera au potentiel du drain de l'autre; et dans le cas des circuits intégrés, le substrat leur est commun. Il s'ensuit qu'une polarisation inverse est alors appliquée entre source ou canal et substrat de l'un des MOS, ce qui modifie ses conditions de fonctionnement.

Cette polarisation accroît la densité des charges fixes Q_B dans la couche de déplétion, ce qui a alors pour effet de réduire le nombre de charges mobiles, et par conséquent d'augmenter la valeur de la tension de seuil. Ainsi, si l'on appelle :

- $Q_B(V)$ — $Q_B(0)$ l'accroissement des charges dans la zone de déplétion;
 - $V_T(V)$ la nouvelle tension de seuil V_{TH} pour une polarisation inverse (V),
- on peut écrire :

$$V_T(V) = V_{TH}(0) - \frac{Q_B(V) - Q_B(0)}{C_O}$$

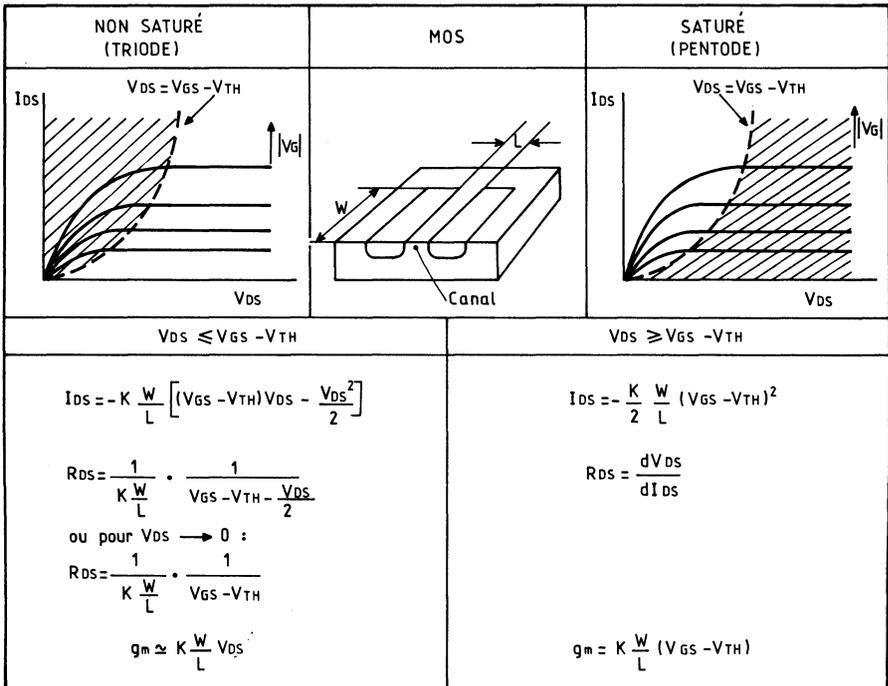


Fig. 12. — Résumé des caractéristiques du MOS en régime triode et pentode.

Fig. 13. — Courbes $I_{DS} = F(V_{DS})$ pour les théories de premier et second ordre.

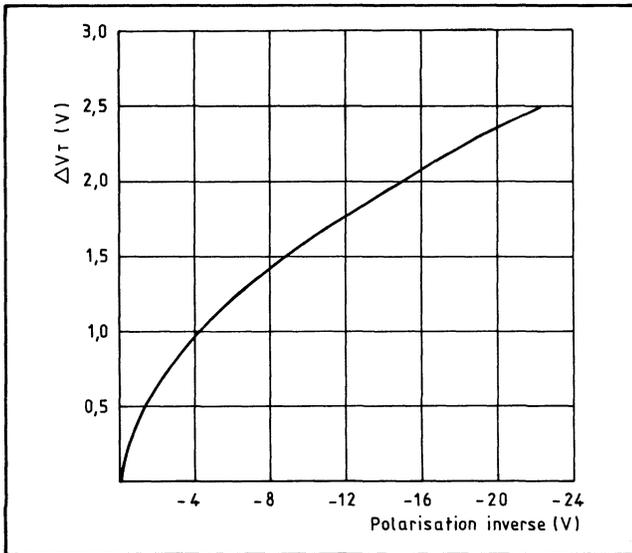
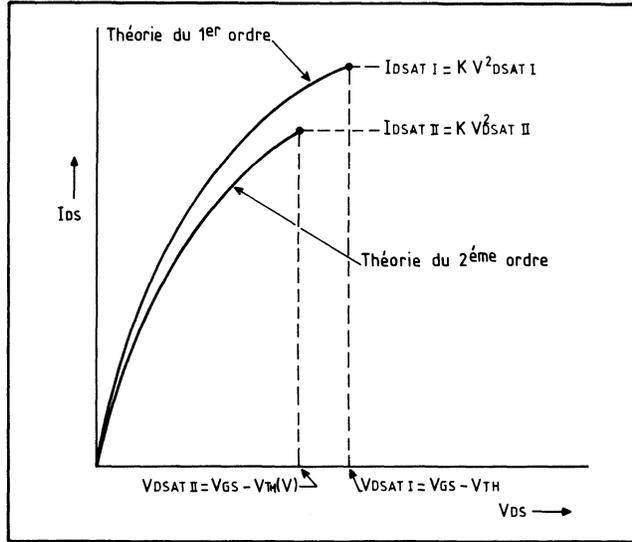


Fig. 14. — Accroissement de la tension de seuil en fonction de la polarisation inverse source à substrat.

A cet accroissement de la tension de seuil, appelé *effet de substrat*, correspond une réduction du courant de drain, calculé très simplement en introduisant dans les formules précédentes la nouvelle valeur V_T (V) au lieu de V_{TH} . Les résultats en sont illustrés par les courbes de la figure 13. La figure 14 montre, elle, comment la tension de seuil s'accroît en fonction de la tension de polarisation inverse source (ou canal) à substrat, pour un niveau de dopage de $N_D = 10^{15}$ at/cm².

Théorie du 3^e ordre

On pourrait encore faire intervenir la longueur réelle du canal en fonction de l'importance de la zone de déplétion qui dépend de la tension de drain. On aboutirait ainsi à une théorie du 3^e ordre, encore plus proche de la réalité, mais dont le développement n'est guère utile ici.

Le facteur de qualité

On a encore pu définir un facteur de qualité, noté ω_0 , qui est le rapport de la pente à la capacité de gate :

$$\omega_0 = \frac{g_m}{C_{OX}}$$

Enfin, et par analogie avec les anciens tubes, un coefficient μ qui est le produit $g_m R_{DS}$ apparaît parfois :

$$\mu = \frac{dV_{DS}}{dV_{GS}} = R_{DS} g_m$$

3. QUELQUES NOTIONS COMPLÉMENTAIRES

Le MOS parasite

On a vu que le transistor MOS, de par sa structure, est auto-isolé du substrat à la condition que ce dernier soit porté à la tension la plus positive. Par conséquent,

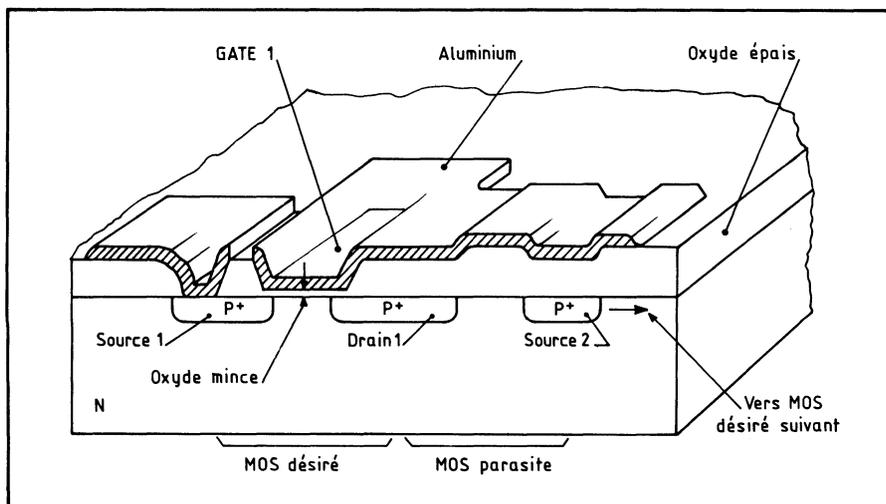
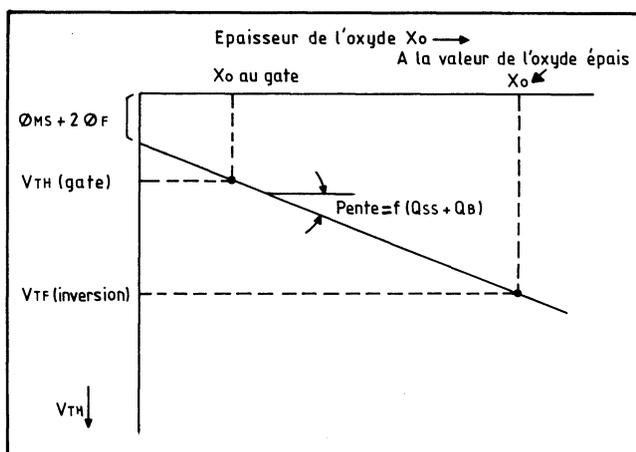


Fig. 15. — Le MOS parasite.

deux MOS voisins implantés sur un même substrat sont également isolés entre eux.

Dans ce cas, pourtant, on peut constater l'apparition de transistors parasites sous certaines conditions. Considérons en effet la figure 15, où un MOS complet a été dessiné, avec une fraction du MOS voisin; entre les deux et au-dessus se trouve une couche épaisse d'oxyde, de l'ordre de $1,5 \mu\text{m}$ (contre 150 nm pour l'oxyde de gate), recouverte d'une bande métallique (aluminium) servant de connexion; pour simplifier le schéma, les prises de contact n'ont pas été toutes dessinées.

Fig. 16. — La tension d'inversion en fonction de l'épaisseur d'oxyde.



On peut considérer que le drain 1 du MOS 1 et la source 2 du MOS 2 constituent un transistor MOS dont le canal est commandé par la tension de gate 1 à travers une couche épaisse d'oxyde. Pour éviter son influence, il faudra que la tension au gate 1 n'atteigne jamais la valeur de la tension de seuil de ce MOS parasite; cette dernière est calculée de la même façon que V_{TH} :

$$V_{TH} = -\frac{Q_{SS}}{C_O} - \frac{Q_B}{C_O} + \Phi_{MS} + 2\Phi_F$$

Soit, avec les mêmes valeurs prises dans les exemples précédents mais avec, cette fois, la capacité par unité de surface divisée par 10 puisque l'épaisseur du diélectrique a été multiplié par 10 :

$$\begin{aligned} V_{TH} &= -10,5 - 14 - 0,35 - 0,5 \\ &= -25 \text{ V} \end{aligned}$$

Cette tension est appelée *tension d'inversion* et désignée par V_{TF} .

La figure 16 montre l'évolution linéaire de V_{TF} en fonction de l'épaisseur de l'oxyde épais. Pratiquement celui-ci mesure au moins $1,5 \mu\text{m}$, toujours pour le MOS fondamental qui sert de base à cette étude, car on verra que d'autres technologies autorisent d'autres valeurs.

La résistance d'entrée du MOS

L'électrode de commande du MOS étant le gate, isolé électriquement du substrat, la résistance d'entrée (en continu) du transistor est énorme. En effet, la

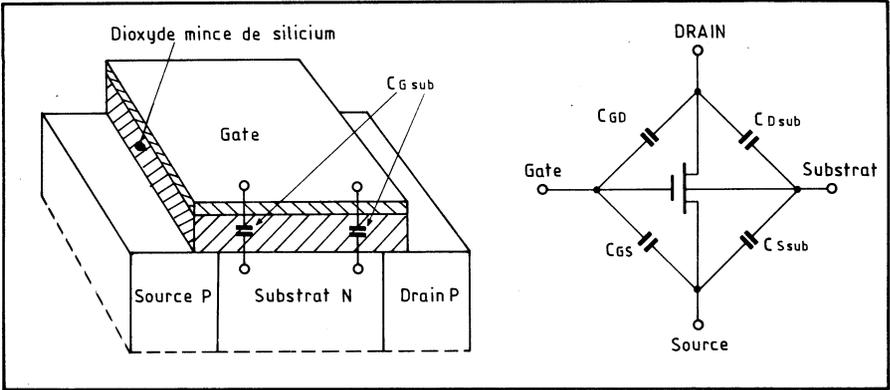


Fig. 17. — Les capacités de structure du MOS.

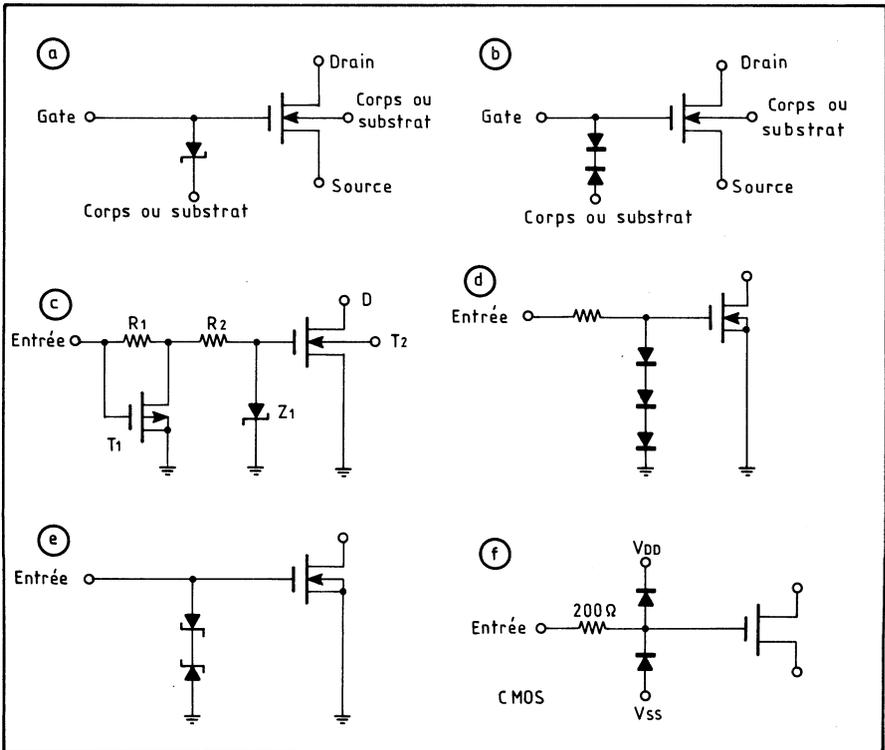


Fig. 18. — Quelques schémas de protection des entrées.

résistivité de l'oxyde thermique de gate est de l'ordre de $10^{17} \Omega \cdot \text{cm}$, pour une rigidité diélectrique de quelque $600 \text{ V}/\mu\text{m}$.

Si l'on suppose que l'épaisseur de silice est de 100 à 200 nm, on trouve une résistance d'entrée de l'ordre, ou supérieure au téraohm ($1 \text{ T}\Omega = 10^{12} \Omega$).

La résistance au claquage est, elle, de l'ordre de 80 V ou supérieure. En fait, ces valeurs dépendent pour beaucoup de la qualité de l'oxyde, et plus précisément de sa « propreté ». C'est là l'une des difficultés de la fabrication des MOS.

Les capacités de structure

Outre la capacité gate-substrat, le MOS comporte encore un certain nombre de capacités (fig. 17) :

- *Gate-drain*, notée C_{GD} ;
- *Gate-source* (C_{GS});
- *Drain-substrat* ($C_{G_{\text{subs}}}$);
- *Source-substrat* ($C_{S_{\text{subs}}}$).

Lorsque la source et le substrat sont reliés au même potentiel, on peut confondre dans C_{GS} les capacités gate-source et gate-substrat. Si tel n'est pas le cas, il convient de séparer ces deux notions.

Ces capacités interviennent de plusieurs façons :

1. *Capacité d'entrée* : elle est, en réalité, la somme des capacités gate-substrat, gate-source, et gate-drain (en série avec drain-substrat, selon qu'on les considère). Il faut encore lui ajouter la capacité des connexions d'accès aux électrodes. Une valeur type de capacité d'entrée, pour un MOS individuel, est ainsi de l'ordre de 4 pF.

2. *Régime dynamique* : cette capacité gate-substrat, parce qu'elle se trouve aux bornes d'une résistance ohmique très élevée, peut jouer le rôle d'un *condensateur-mémoire*. Une fois chargée, elle le reste pendant un temps long, par rapport aux périodes utilisées en électronique : plusieurs millisecondes (2 à 5), par exemple. Elle interviendra alors dans un mode de fonctionnement original, inconnu des bipolaires : le fonctionnement en *régime dynamique*.

3. *En commutation* : on constate que la capacité gate-drain est montée en contre-réaction. Elle intervient ainsi en limitant les performances du MOS et l'on tendra toujours à la minimiser.

La protection des entrées

Si l'on suppose que le diélectrique de gate ne peut supporter que 80 V, on peut imaginer qu'une tension de 100 V qui lui serait appliquée provoquerait un phénomène destructif irréversible.

La capacité de gate du MOS étant de l'ordre de 4 pF, cette tension d'entrée pourra être atteinte à partir d'une charge Q de :

$$Q = CV = 4 \cdot 10^{-10} \text{ coulomb}$$

puisque la résistance d'entrée du MOS en continu est énorme. Or, cette charge est, elle, excessivement faible; elle correspondrait, par exemple, à un courant d'une fraction de nano-ampère (0,4 nA) appliqué pendant une seconde!

On conçoit dès lors que de tels courants ou charges soient à redouter; on est ainsi obligé, en travaillant avec des MOS, de prendre les plus grandes précautions

contre les courants accidentels (fer à souder mal isolé...) ou les charges statiques (l'opérateur porte des vêtements en fibres synthétiques...). On recommande généralement aux opérateurs travaillant sur des circuits intégrés MOS de se mettre à la masse à l'aide d'un bracelet métallique autour du poignet, et cela bien que les entrées des CI soient généralement protégées par le fabricant.

Cette protection revêt plusieurs aspects et peut être assurées par une ou plusieurs diodes (fig. 18), intégrées par le fabricant sur chaque entrée, éventuellement associées à un MOS supplémentaire.

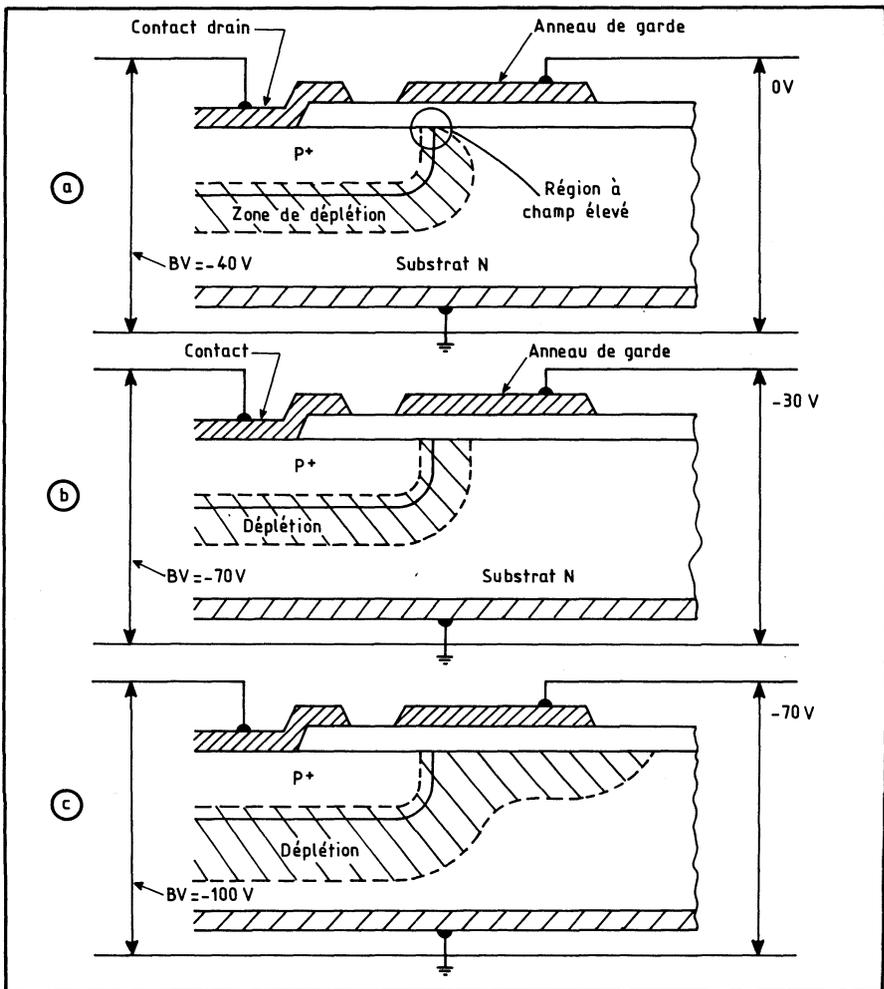


Fig. 19. — Le rôle des anneaux de garde.

D'autres précautions consistent à :

- stocker les circuits intégrés avec toutes les broches en court-circuit, via un support spécial ou une mousse conductrice;
- relier les broches inutilisées à la masse;
- relier le fer à souder à la masse;
- travailler sur des tables recouvertes de mousse conductrice.

L'expérience montre que le non-respect de ces règles ne mène pas obligatoirement à la destruction des circuits, mais qu'à l'inverse leur respect reste encore le meilleur garant de leur bon état de fonctionnement. On reviendra sur ce point dans le chapitre consacré aux CMOS.

Autres limites en tension

La couche de déplétion qui ceinture le MOS devient très étroite à la jonction entre le drain et la surface du substrat. De ce fait, il existe là un champ élevé qui limite la tension que l'on peut appliquer au drain.

Une façon de tourner cette difficulté consiste à édifier un anneau de garde qui, en fonction de sa polarisation, fera reculer cette limite (fig. 19).

CHAPITRE III

LES DIVERS TYPES DE MOS

On a analysé le fonctionnement du MOS fondamental, c'est-à-dire à canal P et à enrichissement. Pratiquement, on connaît de nombreux types de MOS :

- *MOS à canal P, à enrichissement.*
- *MOS à canal N, à enrichissement, dont les performances sont supérieures.*
- *MOS à canal N ou P, mais à déplétion (appauvrissement) : ils sont conducteurs en l'absence de tension de commande.*
- *MOS tétrode, ou à double gate.*
- *MOS à gate flottant, servant de mémoire.*
- *MNOS, destiné également à servir de mémoire mais avec effacement et inscription électriques.*

1. INTRODUCTION

La structure fondamentale présentée antérieurement était celle d'un MOS à canal P, à enrichissement. Or, ceux à canal N offrent des propriétés bien plus intéressantes :

1. Dans les MOS à canal P, la conduction est assurée par des porteurs positifs, des trous, dont la mobilité est de 2 à 3 fois inférieure à celle des porteurs négatifs, les électrons.

Ainsi, sous un champ électrique tel qu'appliqué aux MOS, la mobilité des trous est de l'ordre de 200 cm^2 par volt-seconde, celle des électrons de $400 \text{ cm}^2/(\text{Vs})$. Un élément à canal N aura donc une résistance *on* (R_{ON}) de canal deux fois moindre que celle de son équivalent à canal P de même géométrie.

2. A caractéristiques égales, la surface occupée par un MOS à canal N est la moitié de celle occupée par un MOS à canal P. La densité d'intégration est donc fortement accrue.

3. Pour ces deux raisons, la vitesse de travail est très sensiblement accrue. En effet, cette vitesse dépend de la constante de temps RC, or R_{ON} est réduit, d'une part, et C qui dépend de la géométrie, est réduit également.

On a cependant commencé par l'intégration des PMOS pour la raison suivante. Les contaminants en fabrication, qui sont la bête noire des producteurs de CI, sont souvent chargés positivement, et des ions positifs parasites tendent à s'accumuler à l'interface oxyde-silicium; cette charge (Q_o) provoque un décalage de la tension de seuil qui tend à débloquent le MOS à enrichissement s'il est à canal N.

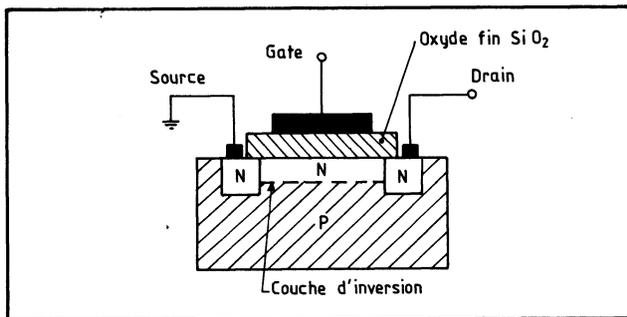


Fig. 1. — MOS à canal N.

D'autre part, la charge fixe Q_{ss} , positive, à l'interface oxyde-silicium, résulte directement du processus de fabrication et décale elle aussi la tension de seuil.

Ces charges existent également dans les MOS à canal P, mais là, les ions $+$ sont attirés vers l'interface aluminium-oxyde par la charge négative de gate et n'affectent guère la tension de seuil; ils ont plutôt tendance à l'accroître sensiblement, mais ce défaut est moins gênant que le phénomène inverse qui débloquent un MOS qu'on souhaiterait bloqué.

On pourrait augmenter le niveau de dopage du MOS à canal N pour compenser ces décalages, mais d'autres considérations montrent que cela ne semble applicable qu'aux MOS complémentaires (CMOS), où l'on peut connecter la source des éléments à canal N à la masse.

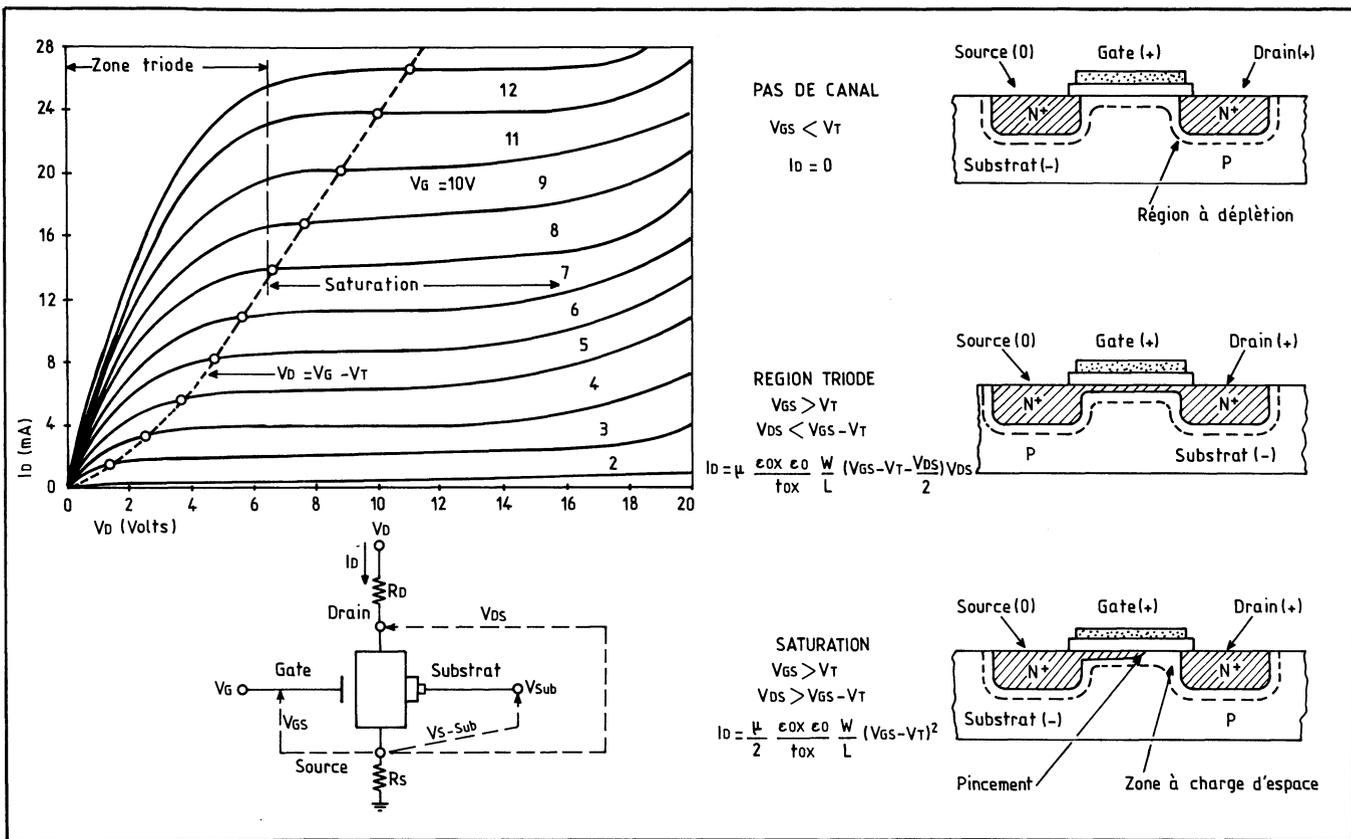
C'est donc bien parce qu'ils sont plus faciles à produire que les MOS à canal P ont été très largement intégrés. Cependant, cette situation est actuellement en train de s'inverser, du fait d'une meilleure connaissance des processus physiques et de la maîtrise de la fabrication, au point que le marché des MOS bascule au bénéfice des NMOS.

2. LE MOS A CANAL N

Fonctionnement et caractéristiques

Dans les NMOS, on a tout simplement inversé les polarisations du substrat, des diffusions, et par suite de l'alimentation qui devient positive (comme pour la TTL bipolaire), comme le montre la figure 1.

Fig. 2. — Caractéristiques typiques du NMOS.



Dès 1971, les MOS à canal N ont commencé à être commercialisés. La société *IBM* les adoptait très rapidement sur ses ordinateurs 370/158 et 370/168, sous forme de mémoires, et étudiait alors des NMOS à couche d'oxyde de 100, puis 50 nm qui vont servir d'exemple ici (1).

Avec ces éléments, on retrouve les caractéristiques typiques des MOS et les situations suivantes (fig. 2) :

1. Si V_{GS} est inférieur à V_{TH} , le MOS est bloqué. Pour une forte différence de tension, il y a accumulation de porteurs P sous le gate; si la différence diminue, cette accumulation fait place à une couche de déplétion.

2. Lorsque V_{GS} est supérieur à V_{TH} , on retrouve les conditions de fonctionnement triode et à saturation.

A saturation, le canal se pince près du drain car la tension de gate est désormais insuffisante pour le rétablir. Cependant, un courant drain continue à se manifester à travers la charge d'espace de la zone pincée. Si l'on accroît encore V_{DS} , la zone pincée s'élargit mais le courant drain reste pratiquement constant.

Le réseau de courbes relevé (fig. 2 a), correspond au NMOS à oxyde de 50 nm d'épaisseur, et canal de longueur $L = 4,6 \mu\text{m}$, de largeur $W = 100 \mu\text{m}$, et pour une polarisation inverse de substrat $V_{\text{sub}} = -7 \text{ V}$. La mobilité des électrons, généralement comprise entre 400 et 500 $\text{cm}^2 \text{V}^{-1}\text{s}^{-1}$, atteint ici 510 $\text{cm}^2 \text{V}^{-1}\text{s}^{-1}$.

Dans les formules donnant le courant drain I_D , on retrouve à nouveau :

ϵ_{OX} constante diélectrique relative de l'oxyde de gate, soit 3,9;

ϵ_0 constante diélectrique du vide = $8,85 \cdot 10^{-14} \text{ F/cm}$;

t_{OX} épaisseur du diélectrique de gate;

μ mobilité des porteurs, donc ici μ_N .

On obtient :

$$I_D (\text{triode}) = \mu \frac{\epsilon_{OX} \epsilon_0}{t_{OX}} \frac{W}{L} \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2} \right) V_{DS}$$

$$I_D (\text{saturation}) = \frac{\mu}{2} \frac{\epsilon_{OX} \epsilon_0}{t_{OX}} \frac{W}{L} (V_{GS} - V_{TH})^2$$

La transconductance est de l'ordre de :

$$g_m (\text{élément } 50 \text{ nm}) = 36 \mu\text{mhos/V}$$

$$g_m (\text{élément } 100 \text{ nm}) = 18 \mu\text{mhos/V}$$

Le substrat étant polarisé, on constate que la tension de seuil se déplace en raison de ce que l'on appelle l'*effet de substrat* comme le montre la figure 3 donnant la courbe de V_{TH} en fonction de la différence de tension $V_{\text{source}} - V_{\text{substrat}}$. Cette tension de seuil est calculée de la façon suivante (toujours selon *IBM*) :

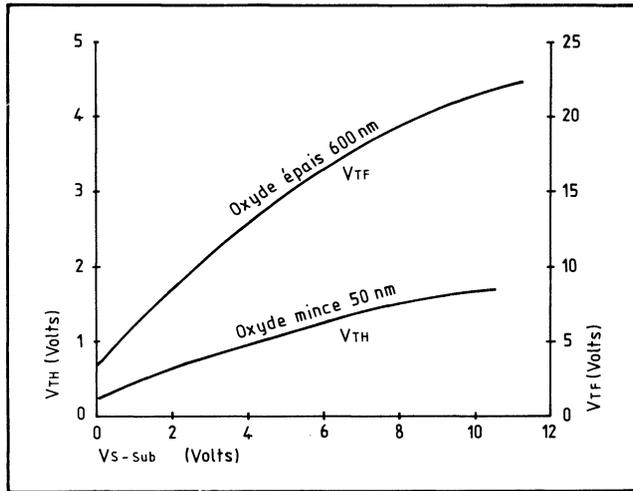
$$V_{TH} = \frac{t_{OX}}{\epsilon_0 \epsilon_{OX}} [-Q_{SS} - Q_B + \sqrt{2 \epsilon_0 \epsilon_{Si} q N_A (V_{S-\text{Subs}} + W_S)}] \\ + \Phi_{MS} + W_S + \Delta V_{DT}$$

1. Une importante étude *IBM* portant sur ce thème a été publiée dans *IBM Journal* septembre 1973, par D. L. CRITCHLOW, R. H. DENNARD et S. E. SCHUSTER : *Design and characteristics of n-channel insulated-gate FET*.

où :

- V_{TH} est la tension de seuil minimale;
- ϵ_{Si} est la constante diélectrique relative du silicium, soit 12;
- q est la charge de l'électron, soit $1,6 \cdot 10^{-19}$ coulomb;
- Q_{SS} est la charge fixe à l'interface Si-SiO₂, dans l'oxyde;
- Q_B est la charge fixe, à l'interface, zone de déplétion;
- N_a est le niveau de concentration en impuretés du substrat;
- V_{S-sub} est la différence des tensions source-substrat;

Fig. 3. — Les tensions de seuil et d'inversion du NMOS en fonction de la polarisation inverse du substrat.



- W_s est la tension aux bornes de la couche de déplétion, à l'établissement de la conduction et en l'absence de polarisation du substrat, soit 0,75 V pour du silicium P de 2,3 Ω/cm;
- Φ_{MS} la différence des niveaux Fermi de l'aluminium et du silicium, soit -0,8 V pour du silicium P;
- ΔV_{DT} le décalage de V_{TH} dû à l'effet de déplétion du bore.

Les courbes de la figure ont été tracées pour deux épaisseurs d'oxyde et montrent donc comment évoluent V_{TH} mais aussi la tension d'inversion V_{TF} correspondant au MOS parasite, au-dessus de l'oxyde épais supposé de 600 nm. On a fait :

$$Q_B + Q_{SS} = 0,5 \cdot 10^{11} \text{ charges/cm}^2$$

avec $N_a = 5 \cdot 10^{15} \text{ cm}^{-3}$ et $\Delta V_{DT} = -0,07 \text{ V}$.

De ces courbes, on peut tirer une conclusion importante : la tension de seuil V_{TH} étant très faible, on peut l'augmenter en polarisant négativement le substrat ; ce faisant, on améliore également la tension d'inversion V_{TF} de l'oxyde épais, et le rapport V_{TF}/V_{TH} devient plus favorable. On admet, en pratique, que le rapport V_{TF}/V_{TH} doit être de l'ordre de 10.

Les capacités de structure

Tout comme le PMOS, le NMOS est doté de capacités parasites parmi lesquelles (fig. 4 a) :

- la capacité *gate-canal* ;
- la capacité *gate-source* ;
- la capacité *gate-drain*.

L'ensemble de ces 3 capacités est généralement regroupé sous le nom de capacité *gate-substrat*, ou capacité de structure (ou *intrinsèque*), et intervient dans le fonctionnement *en dynamique* du MOS. S'y ajoutent :

- la capacité *source-substrat* ;
- la capacité *drain-substrat*.

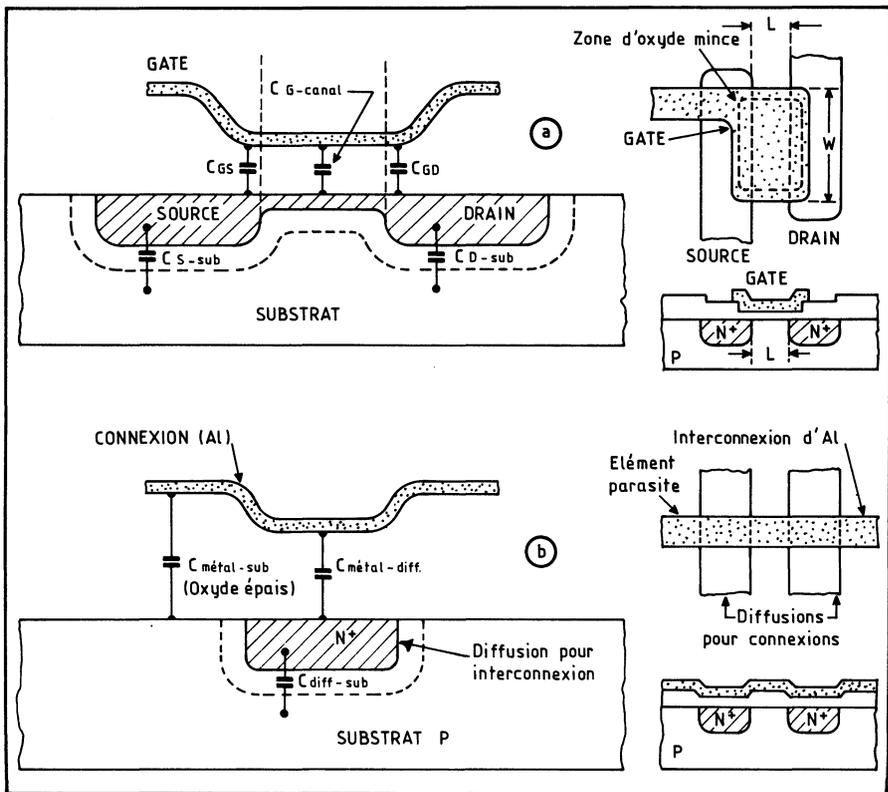


Fig. 4. — Les capacités de structure du NMOS.

Et enfin, les capacités dues aux connexions métalliques (en b) :

- *conducteur-diffusion pour connexion* ;
- *conducteur-substrat*.

Exemples de valeurs de courant drain

Si l'on tient compte des résistances séries R_D et R_S (fig. 2) dues principalement à la résistance superficielle des diffusions, on peut élaborer un réseau pratique de courbes donnant le courant drain en fonction de la tension. Dans les figures 5, plus exactement, on a utilisé en échelle verticale le rapport $I_D/(W/L)$, et ce pour les NMOS à épaisseur d'oxyde de gate de 50 et 100 nm.

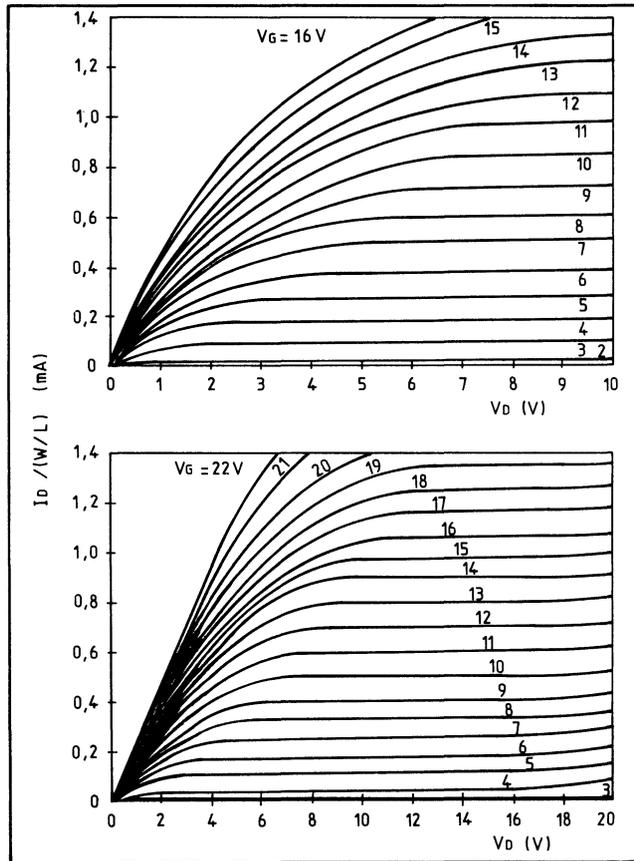


Fig. 5. — Courbes typiques du NMOS.

Précisons, à ce propos, que L est la valeur réelle, effective de la longueur du canal, limitée par les diffusions, et non l'espace réservé au canal si les diffusions ne s'étaient pas horizontalement. Ainsi, pour le MOS de 50 nm, la longueur théorique était de 9 μm pour un L effectif de 4,5 μm . Pour le MOS de 100 nm, la longueur réservée était de 12 μm pour un L réel de 7 μm environ. De la même façon, les courbes de I_D qui seraient dues au MOS parasite sont tracées figure 6, dans les conditions définies précédemment.

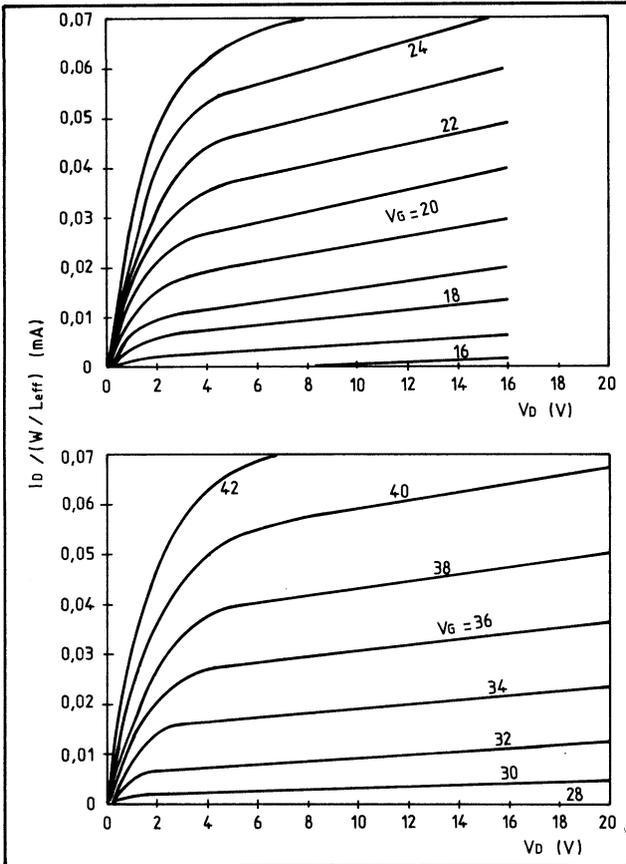


Fig. 6. — Courbes du NMOS résultant d'une tension d'inversion appliquée sur l'oxyde épais.

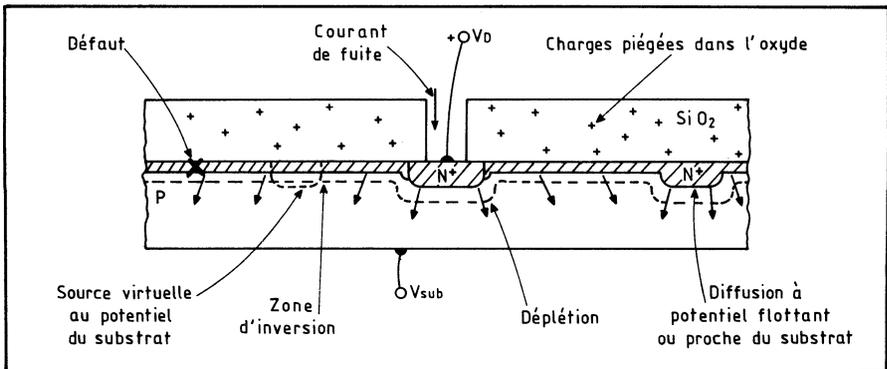


Fig. 7 — La naissance des courants de fuite en surface.

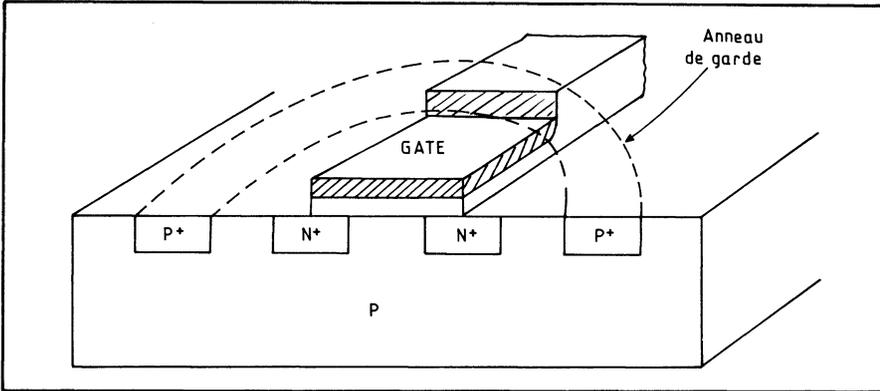


Fig. 8. — Les anneaux de garde interdisent les courants de fuite.

Les courants de fuite

Comme on l'a évoqué plus haut, les charges positives fixes piégées dans l'oxyde sont plus redoutables avec les NMOS qu'avec les PMOS. En effet, et en raison de la faible tension de seuil inhérente aux NMOS, si ces charges s'accroissent elles peuvent inverser l'état de la surface du silicium et donner naissance à un courant allant des diffusions N+ vers la couche d'inversion ainsi créée. Poursuivant sa route, ce courant peut atteindre soit d'autres diffusions N, soit le substrat via la zone de déplétion. Enfin, il est évident que le moindre défaut favorisera ce courant de fuite.

Pour pallier cette difficulté, on diffuse classiquement des anneaux de garde P+, soit autour du MOS (fig. 8), soit autour des diffusions N+ de connexions. Ou

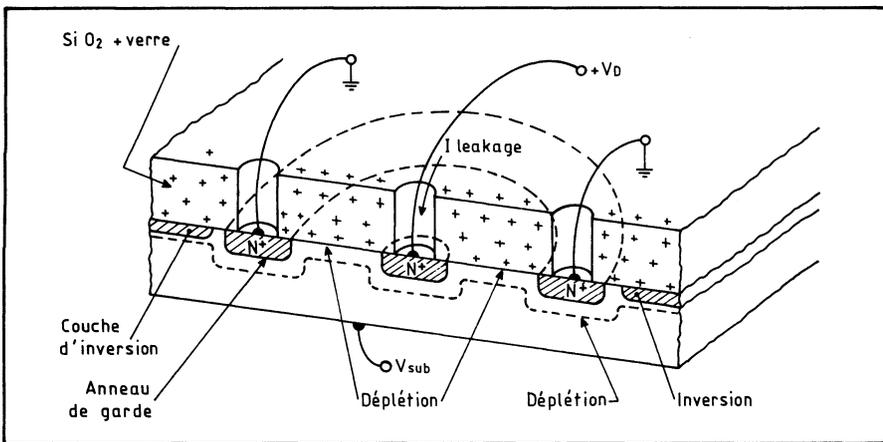


Fig. 9. — Anneau de garde par diffusion N+.

encore, et selon une technologie utilisée par *IBM*, on diffuse un anneau de garde N^+ pour éviter l'opération supplémentaire que constitue la diffusion P . Dans ce dernier cas, en effet, on peut considérer l'anneau de garde comme la source d'un MOS. Or, si le substrat est lui-même polarisé en inverse, on peut considérer que la « tension de seuil » se sera accrue en raison de l'effet de substrat, fermant le chemin au courant de fuite (fig. 9).

3. LES MOS A DÉPLÉTION

Dans les MOS à enrichissement, qu'ils soient à canal N ou à canal P , on *enrichit* le canal en porteurs minoritaires pour les rendre conducteurs. Ainsi, le *MOS à enrichissement est bloqué en l'absence de signal de commande*. A l'inverse, les MOS à appauvrissement, ou à déplétion, sont *conducteurs en l'absence de signal de commande*. En effet, un canal est créé en cours de fabrication du transistor (ce qui exige une opération de diffusion supplémentaire, notons-le au passage), et il existe en l'absence de la tension V_G .

Les MOS à déplétion sont à canal N ou à canal P . Dans la figure 10 est représenté un MOS à canal N , à déplétion, vu en coupe, comparé à un MOS à enrichissement à canal N également.

Pour bloquer ce MOS à déplétion, conducteur en l'absence de signal de gate, il faut appliquer une tension V_G négative qui chasse les porteurs majoritaires du substrat sous-jacent, donc appauvrit le canal jusqu'à le faire disparaître.

Les courbes I_D fonction de V_{GS} de la figure 11 comparent le comportement de MOS à enrichissement et à déplétion. Ces courbes sont simplement translatées, et l'on pourrait obtenir toute caractéristique intermédiaire en agissant sur le dopage du canal.

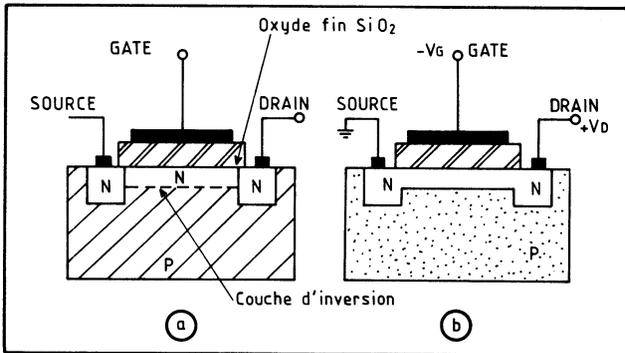


Fig. 10. — Structure d'un MOS à enrichissement à canal N (a), ou à déplétion (b).

En effet, le canal n'étant pas dopé, on obtient un MOS à enrichissement dont la tension de seuil sera, par exemple, de 4 V (canal N). Si l'on dope le canal avec des impuretés N , cette tension de seuil commence par diminuer et passe par toutes les valeurs intermédiaires pour s'inverser dès lors qu'un canal N est créé. On verra, plus loin, que ce dopage précis est assuré par implantation ionique.

Fig. 11. — Plus on implante le canal et plus la tension de seuil se déplace.

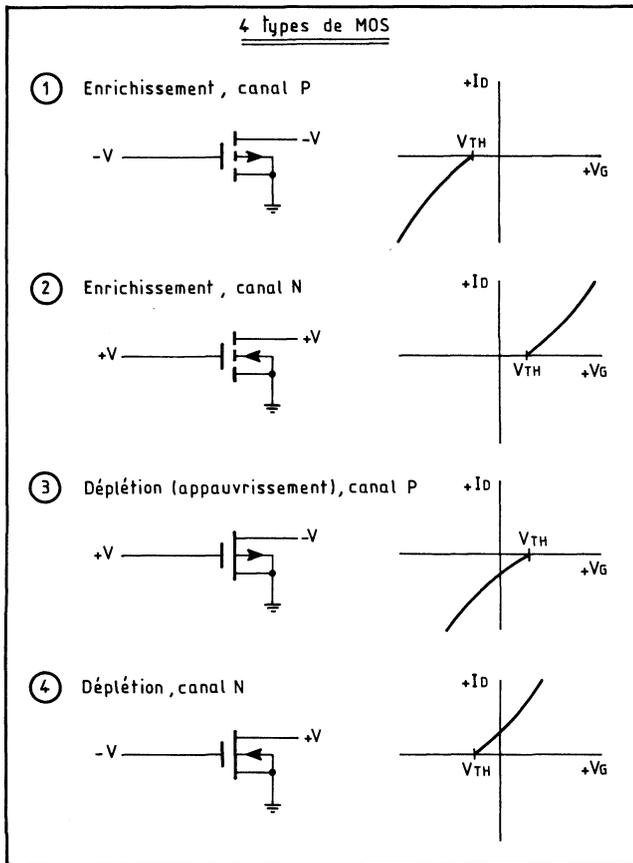
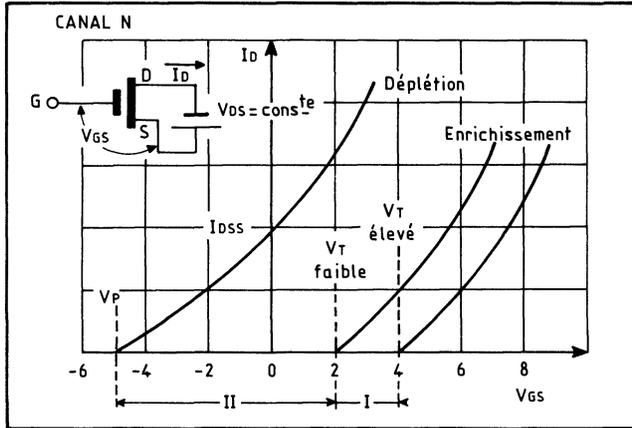


Fig. 12. — Les 4 types de MOS.

Pour éviter toute confusion, on est convenu de représenter les MOS à appauvrissement et à enrichissement, canal N et canal P, comme le montre la figure 12. Les MOS à enrichissement étant bloqués en l'absence de tension de commande, leur canal est interrompu sur le schéma. Ceux à appauvrissement étant conducteurs en l'absence de tension de commande, leur canal est représenté par un trait continu.

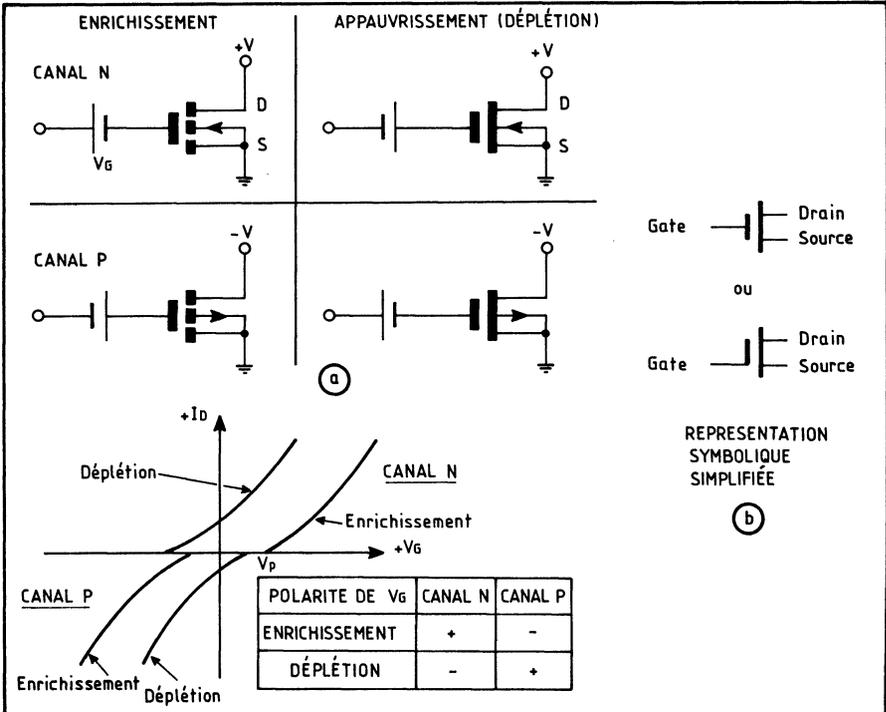


Fig. 13. — Caractéristiques comparées des 4 MOS.

Le sens de la flèche indique la polarité du canal. Un moyen de se souvenir de sa signification consiste à assimiler le canal à la base d'un transistor bipolaire; on a alors :

- flèche vers l'extérieur égal canal P, comme pour les NPN où la base est P;
- flèche vers l'intérieur égal canal N, comme pour les PNP où la base est N.

Cette flèche est portée sur la connexion représentant le substrat, importante à respecter surtout lorsque le substrat n'est pas relié à la source.

Pratiquement, on simplifie généralement les schémas lorsque l'ensemble de ces informations n'est pas indispensable. On revient alors au symbole général du MOS, donné à nouveau figure 13 avec les variantes regroupées et la polarité à appliquer au gate (V_G) pour commander le dispositif.

4. LE MOS TÉTRODE (MOS A DOUBLE GATE)

Deux électrodes de commande peuvent être réalisées sur un même MOS (fig. 14) : le nouvel élément prend alors le nom de *MOS à double gate*, ou *MOS tétrade* en raison de ses 4 électrodes (drain, source et 2 gates). Le MOS « normal » est donc appelé par analogie, *MOS triode*.

On utilise souvent la tétrade MOS dans l'équivalent du montage cascade (fig. 15); ce montage permet de réduire la capacité de réaction et donc d'accéder à des fréquences de travail élevées (plusieurs centaines de mégahertz, sinon 1 à 2 GHz). La structure d'une tétrade MOS avec gates protégés, enfin, est représentée figure 16 (selon *RCA*).

5. LE MOS A GATE FLOTTANT

Une structure originale de MOS, proposée par *Intel* en 1970 pour la réalisation de mémoires mortes reprogrammables par l'utilisateur, est le MOS à gate flottant. Sa structure est identique à celle d'un MOS classique à enrichissement, la différence résidant dans le fait qu'aucune sortie n'a été ménagée pour le gate.

Fig. 14. — Schéma électrique (a) et structure (b) d'un MOS à double gate (encore appelé MOS tétrade).

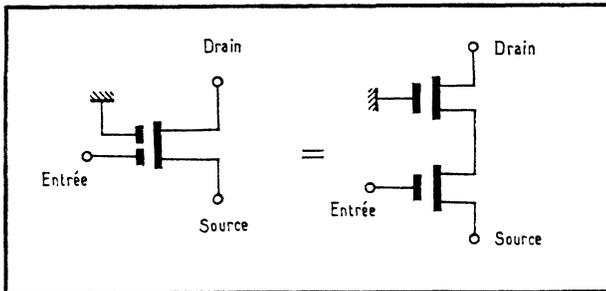
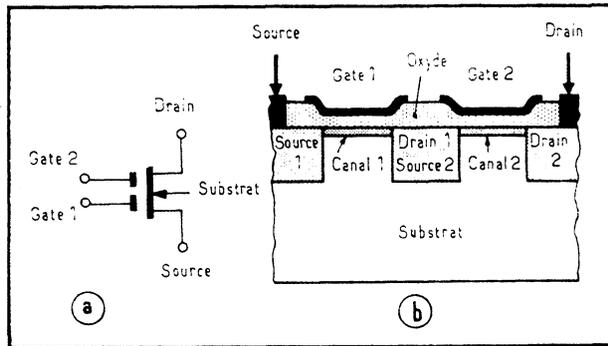


Fig. 15. — Principe d'un montage cascade avec un MOS à double gate (le montage cascade est souvent utilisé en VHF, en télévision par exemple).

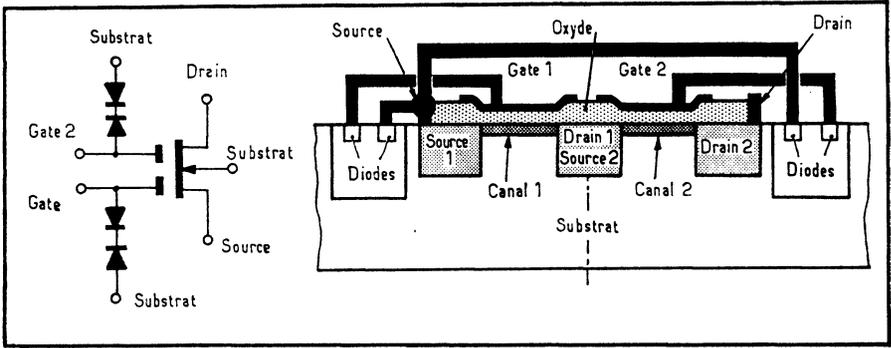


Fig. 16. — Le MOS à double gate est, ici, protégé sur ses entrées par deux paires de diodes montées tête-bêche. En (a), schéma équivalent; en (b), structure physique.

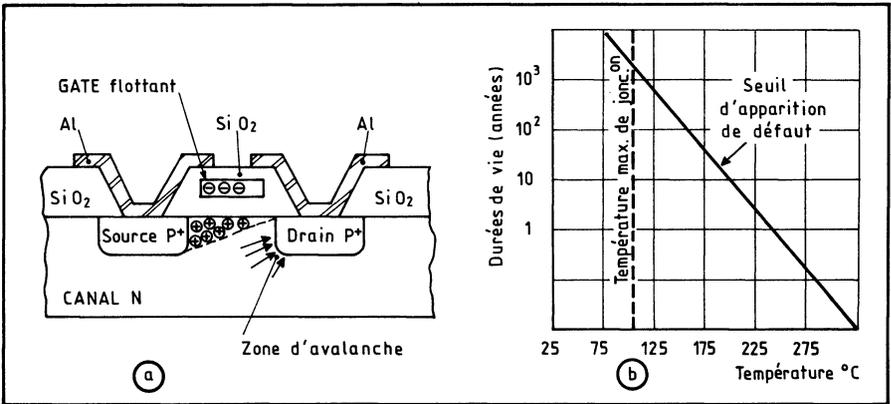


Fig. 17. — Le FAMOS de Intel conserve sa charge pendant plus d'un siècle.

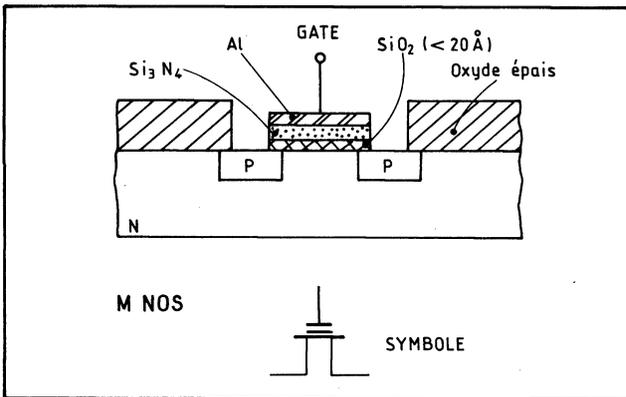


Fig. 18. — Le transistor MNOS.

Celui-ci est donc inaccessible et totalement enrobé et isolé par de l'oxyde (fig. 17).

Entre l'électrode conductrice de gate et le substrat se trouvent environ 100 nm de silice; au-dessus, la couche de silice est dix fois plus épaisse : 1 000 nm. Un tel MOS constitue cependant un élément actif car *le gate peut être chargé à travers les 100 nm de silice, par effet tunnel* : si l'on applique une impulsion fortement négative au drain de ce PMOS, la jonction PN constituée par le drain et le MOS part en avalanche et, par effet tunnel, des électrons sont injectés dans le gate flottant. Leur nombre dépend de l'amplitude et de la durée de l'impulsion. Lorsque celle-ci a disparu, les électrons restent piégés dans le métal du gate sans possibilité de s'échapper : le gate est donc chargé et provoque la naissance d'un canal P dans le substrat. Le transistor est rendu conducteur en permanence.

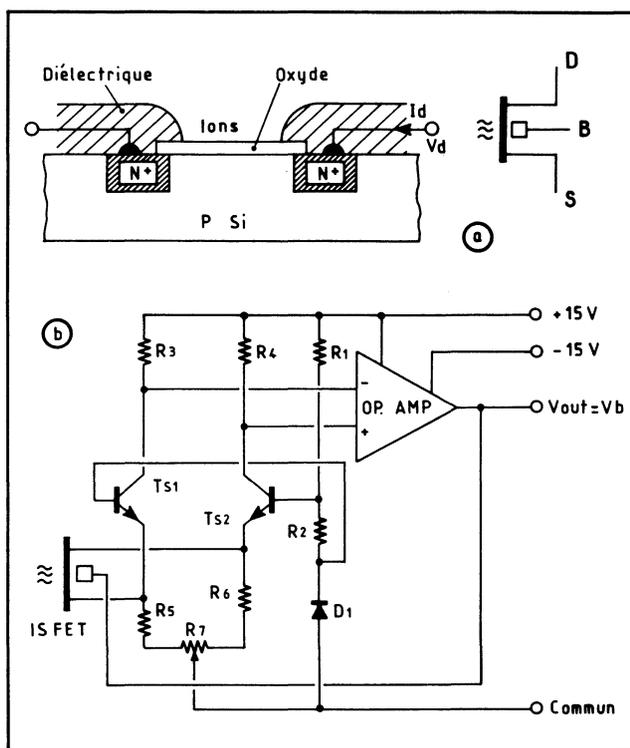


Fig. 19. — Coupe et symbole de l'ISFET, détecteur d'ions (a). Montage d'utilisation (b).

Ce MOS, ou FAMOS, de *floating avalanche injection MOS*, a ainsi mémorisé une information qui ne sera pas détruite, même si l'alimentation est supprimée : on dit que le point-mémoire n'est pas *volatile*.

La charge stockée se conserverait plus d'un siècle dans le gate à la condition de ne pas dépasser 100 °C. On peut cependant l'effacer en soumettant le MOS à un rayonnement ionisant; ce dernier, de l'ultraviolet par exemple, provoque la création d'un courant de fuite capable de décharger le gate. Pratiquement, une source d'UV de 253,7 nm avec une intensité de 10 mW/cm², située à 4 cm du circuit intégré, suffit à l'effacer en moins de 20 minutes; encore faut-il alors que le circuit soit accessible aux UV, c'est-à-dire qu'il soit muni d'une fenêtre transparente.

L'une des premières mémoires à transistors FAMOS, la 1701, à enregistrement électrique et effacement par rayonnement, était ainsi commercialisée dès 1971.

6. LE MNOS

Le transistor MNOS (de *métal, nitrure, oxyde, silicium*) ressemble au MOS à nitrure de silicium (qu'on étudiera dans le chapitre consacré aux technologies). Il est représenté en coupe dans la figure 18 avec son symbole.

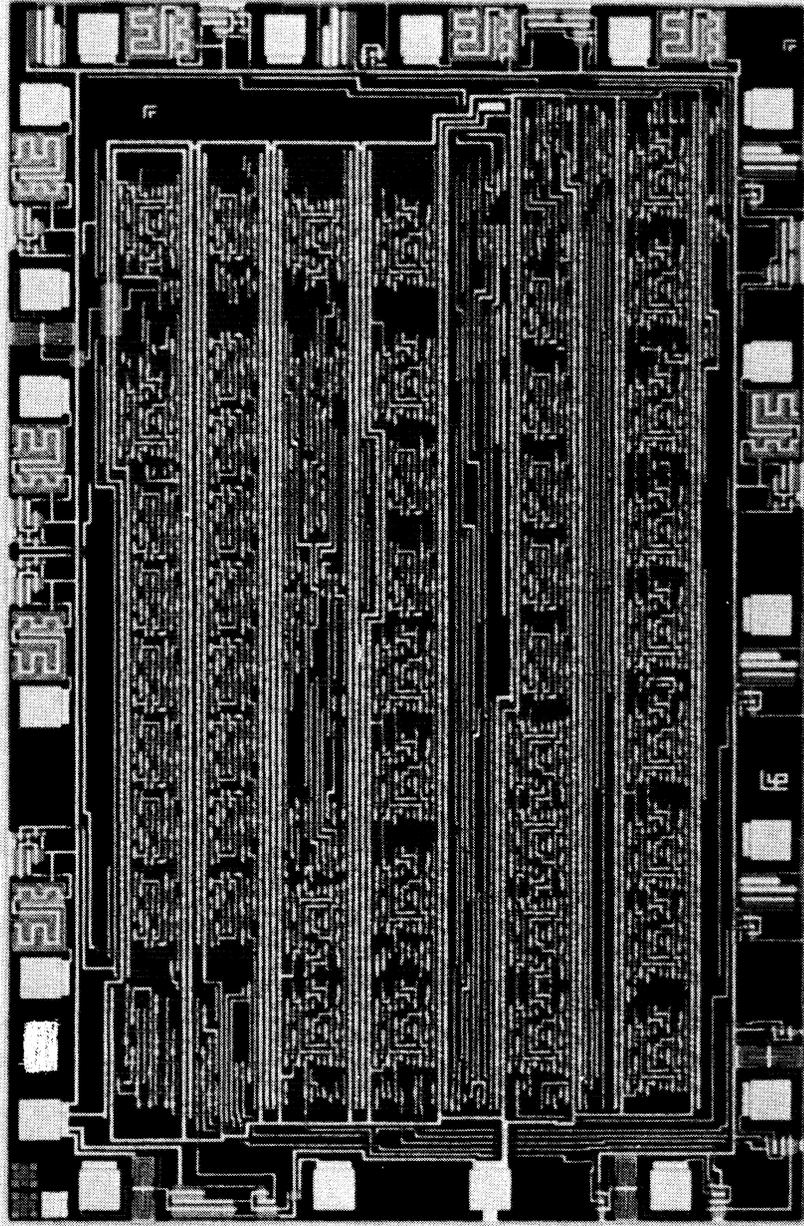
Ici, une couche de nitrure de silicium Si_3N_4 a été déposée sur une couche très mince de silice, de moins de 2 nm d'épaisseur. De ce fait, et si l'on polarise positivement le gate sous des tensions relativement importantes, le drain et la source étant normalement alimentés, des électrons traversent l'oxyde mince par effet tunnel et restent piégés à l'interface oxyde-nitrure. La polarisation initiale disparaissant, cette interface joue le rôle de gate et, étant chargée, maintient le MOS conducteur. A l'inverse, si l'on polarise négativement et fortement le gate, les électrons piégés sont chassés et, toujours par effet tunnel, retraversent la silice en sens inverse.

On dispose ainsi d'un MOS jouant le rôle d'un point-mémoire, mais programmable et effaçable *électriquement* cette fois, sous des tensions de quelques dizaines de volts (30 à 50). Comme avec le FAMOS, la charge acquise se conserverait très longtemps sans perte notable.

Activement étudié, ce MNOS pourrait répondre à un besoin important de l'électronique et servir à réaliser une mémoire RAM non volatile. Ajoutons que plusieurs variantes de ce transistor ont été réalisées en laboratoires; elles portent surtout sur les polarisations à appliquer pour le charger ou le décharger. Les premiers MNOS ont pu ainsi être diffusés dès 1974.

7. LE MOS DÉTECTEUR D'IONS

En supprimant la métallisation de gate, un chercheur hollandais, Piet BERGVELD, a réalisé un MOS détecteur d'ions, présenté au *Congrès Biocapt* (3 novembre 1975) sous le nom de ISFET (pour « *ion sensitive FET* »). Le transistor est utilisé en capteur, la couche d'oxyde jouant le rôle de l'élément sensible aux ions; en fait, ceux-ci constituent le gate du MOS; le corps du transistor sert d'entrée supplémentaire pour appliquer une contre-réaction linéarisant la mesure (fig. 19).



Une structure intégrée CMOS/SOS (sur substrat isolant) fortement grossie. Ce circuit, dû à EFCIS (Grenoble) peut fonctionner jusqu'à 25 MHz.

CHAPITRE IV

LA FABRICATION DES CIRCUITS INTÉGRÉS

1. LA TECHNOLOGIE PLANAR

La technologie planar qui sert fondamentalement de base à la fabrication des circuits intégrés, depuis son invention en 1960 par *Fairchild*, est la suivante.

Préparation du silicium

Le matériau de départ est du silicium (Si) polycristallin; il est purifié de façon à ne contenir au maximum qu'un atome de corps étranger pour 10^{10} atomes de Si, puis transformé en silicium monocristallin par tirage, opération au cours de laquelle on procède à un dopage P ou N.

Doper un matériau, c'est introduire dans son réseau cristallin des impuretés P et N qui lui fournissent des porteurs, trous ou électrons; ceux-ci transforment le semiconducteur à haute résistivité en un conducteur à faible résistivité (quelques ohms-centimètres à quelques dizaines d'ohms-centimètres), polarisé de surcroît.

On dispose désormais d'un lingot cylindrique, dans lequel un méplat est meulé pour fournir une référence au plan cristallographique. En effet, on peut assimiler le silicium à un cristal cubique, qui peut être découpé en tranches selon plusieurs plans, repérés par des indices (fig. 1). Le plan le plus généralement utilisé est le $\langle 111 \rangle$, mais on trouvera également le $\langle 100 \rangle$.

Ce cylindre monocristallin a un diamètre qui s'est constamment accru au cours des années : de 20 mm, on est ainsi passé à 100 mm (4 pouces, exactement) en 1978. Dans le tableau A, on a donné les surfaces correspondantes de silicium et le nombre de circuits qu'on pouvait fabriquer, pour des circuits de 5 mm² et de 60 mm², ces derniers étant parmi les plus importants en 1978. On notera, à ce propos, que le rendement d'une tranche en circuits fonctionnant parfaitement est loin d'être de 100 %!

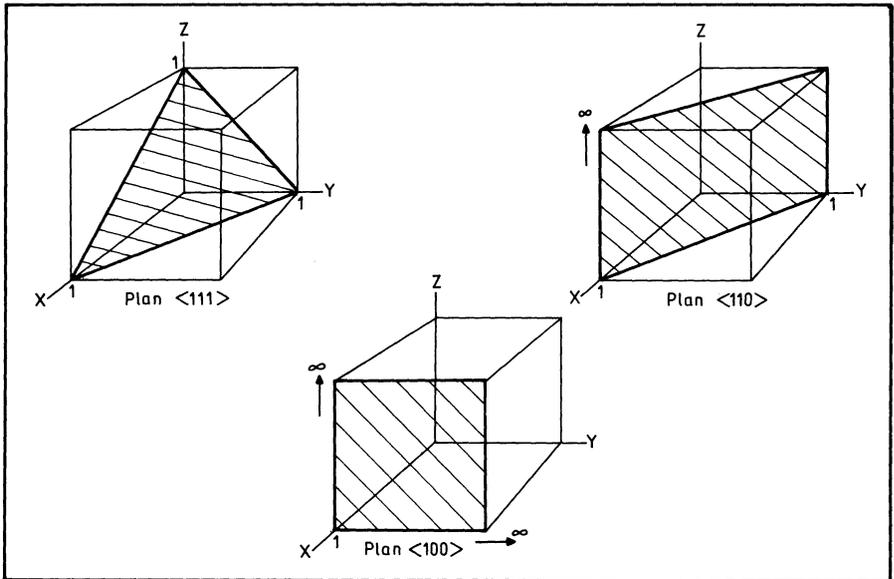


Fig. 1. — Trois plans cristallographiques du silicium.

Chaque tranche a une épaisseur de l'ordre de 0,3 mm; les tranches seront ensuite rodées et polies, à l'aide d'abrasifs et/ou chimiquement. La tranche est souvent désignée par son nom américain : *wafer* (fig. 2 A).

Pour suivre le développement ultérieur des opérations, on va maintenant examiner ce qui se passe dans une petite fraction de la tranche, au niveau de l'intégration d'un unique transistor (fig. 2), à partir de l'opération B.

Oxydation

Les tranches de silicium sont placées dans un four de diffusion, porté à haute température (900 à 1 200 °C). Elles se trouvent sur un porte-échantillon en quartz.

Un agent oxydant qui est de l'oxygène gazeux, ou de la vapeur d'eau, est envoyé sur les tranches et se combine au silicium pour former de la silice, ou dioxyde de silicium SiO_2 .

Photogravure

Pour délimiter les zones qui doivent être dopées, on recourt aux classiques procédés de la photogravure. Pour cela, le schéma électrique du circuit est transformé en dessin d'implantation physique des éléments (zones N, P, isolants, conducteurs, prises de contacts...), décomposé ensuite en schémas partiels relatifs chacun à une, et une seule, opération de fabrication.

Ces dessins partiels à *grande échelle* sont reproduits sur un film transparent garni d'une couche opaque; cette dernière est découpée et ôtée et l'on obtient finalement une vaste feuille transparente garnie de zones opaques qui reproduisent le schéma original. Elle est photographiée avec la réduction appropriée pour amener le tout à l'échelle 1/1.

Tableau A. — Le diamètre des tranches de silicium, leur surface et le nombre de CI

DIAMÈTRE (mm)	SURFACE (mm ²)	ACCROISSEMENT DE LA SURFACE PAR RAPPORT A		NOMBRE DE C.I. PAR TRANCHES	
		Ø=20mm	Ø précédent	C.I. de 5 mm ²	C.I. de 60 mm ²
20	314	---	---	62	5
25	491	1,5	1,5	98	8
30	708	2,2	1,6	140	11
38	1140	3,6	1,6	228	19
50	1960	6,2	1,7	392	32
75	4420	14	2,2	884	73
100	7850	25	1,7	1570	130

Cette photo ne couvrant pas l'ensemble de la tranche de silicium, on la reproduit autant de fois qu'il est nécessaire, côte à côte, pour obtenir un quadrillage parfait qui correspond à la surface du wafer.

Une telle reproduction photographique n'est cependant pas assurée sur un film classique, trop fragile, mais sur un support rigide de verre avec des motifs opaques en chrome. C'est ce qui constitue *le masque*.

Son utilisation est la suivante : sur une tranche de silicium préalablement oxydée sur toute sa surface, on dépose par centrifugation (à l'aide d'une tournette) un vernis photosensible, appelé *photorésist*, durci ensuite par cuisson. Au-dessus, on place le masque et l'on insole, généralement en ultraviolet (en C).

On « développe » ensuite le wafer, en éliminant partiellement la laque, ce qui laisse apparaître l'oxyde de silicium selon des motifs reproduisant fidèlement ceux du masque (en D).

Puis, on attaque l'oxyde nu afin d'avoir accès au silicium. L'oxyde protégé par la laque polymérisée restant après le développement ne peut pas, lui, être attaqué (E). Détruire localement l'oxyde de silicium s'appelle « ouvrir une fenêtre ».

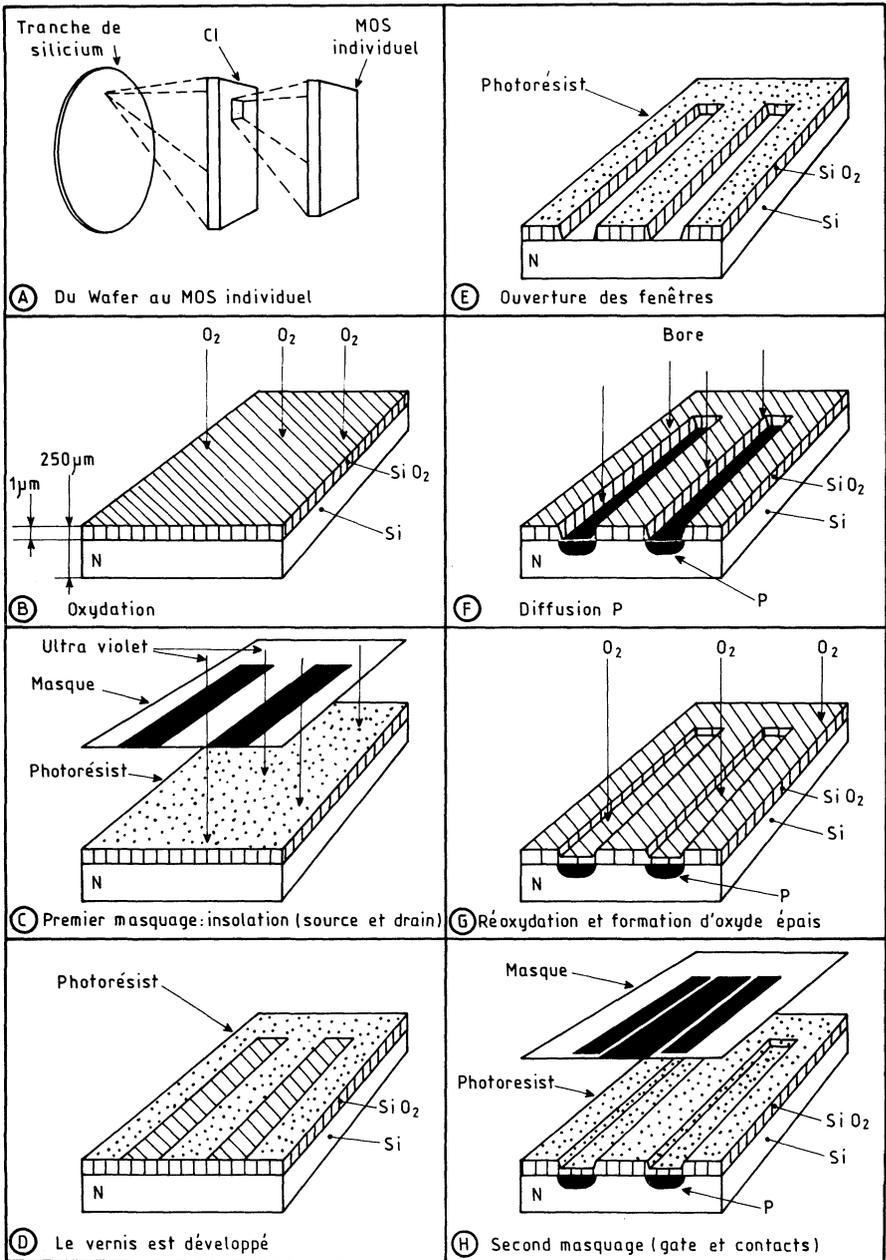


Fig. 2. — Fabrication d'un circuit intégré MOS (suite page ci-contre).

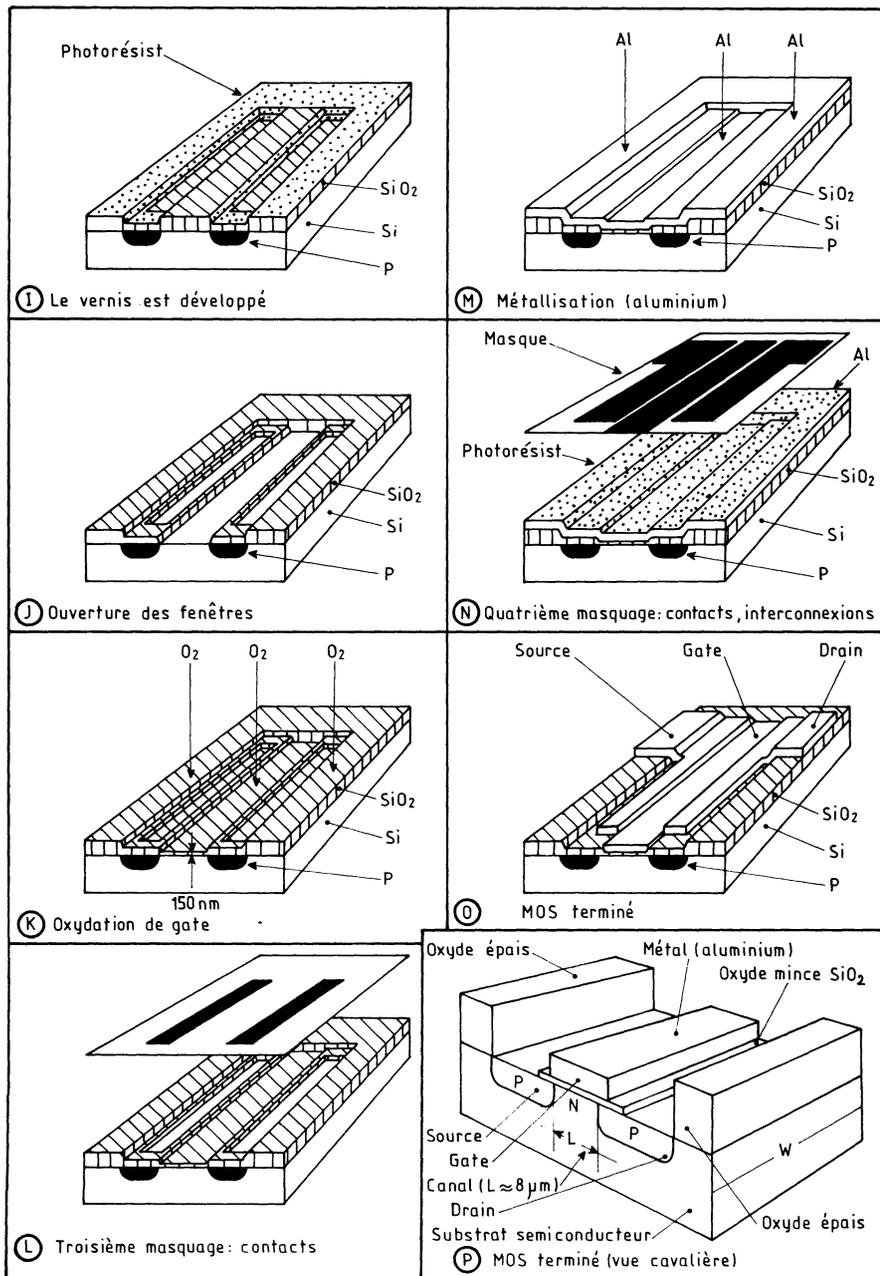


Fig. 2 (suite). — Fin du film de la fabrication d'un MOS.

Le dopage par diffusion gazeuse

C'est à travers ces fenêtres que l'on procède au dopage, par diffusion. C'est une opération à haute température, 1 100 °C environ, assurée dans un four et très délicate à mener : les impuretés véhiculées par un gaz se déposent dans les fenêtres pour doper le silicium (*F*).

Tableau B. — Fabrication fondamentale comparée MOS — bipolaires classiques

MOS	BIPOLAIRES
1 seule diffusion	4 à 6 diffusions
Pas de caissons d'isolement	Nécessité de caissons d'isolement (30% de la surface)
Pas d'épitaxie	Epitaxie
3 à 5 masquages	6 à 8 masquages
38 opérations	130 opérations
2 opérations à >1000°C	10 opérations à > 1000°C
Gain du MOS indépendant diffusion	Gain du bipolaire lié à la diffusion
Structure de l'inverseur plus simple	Inverseur plus complexe
Surface plus faible (donc orientation vers LSI)	Surface plus grande

Cette opération est répétée autant de fois qu'il faut de dopages différents. Le plus grand soin est alors pris pour que les masquages successifs se superposent parfaitement; la précision est de l'ordre du micron.

Le dopage est suivi par une diffusion en profondeur, ou *recuit*, au cours de laquelle les impuretés pénètrent dans le cristal sur une épaisseur de plusieurs micromètres. Au cours de cette opération est assurée une ré-oxydation superficielle (*G*).

Autres étapes

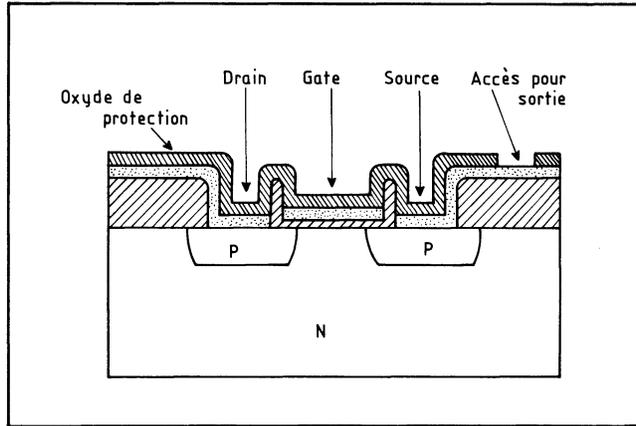
Les étapes suivantes consistent à répéter des opérations semblables :

- (*H*) Second masquage pour la création du gate et des contacts.
- (*I*) Développement du vernis photosensible.
- (*J*) Ouverture de la fenêtre de gate.
- (*K*) Oxydation mince de gate.
- (*L*) 3^e masquage, pour les contacts.
- (*M*) Métallisation de toute la tranche à l'aluminium. Cette opération est assurée par pulvérisation sous vide.
- (*N*) 4^e masquage, délimitant les contacts et les connexions sur le circuit intégré.
- (*O*) et (*P*) Aspect du MOS intégré.

Souvent, enfin, le circuit intégré terminé sera protégé par une couche d'oxyde ou de verre (fig. 3).

Lorsque toutes ces opérations sont terminées, on raye la tranche au diamant pour créer des lignes de cassure qui permettent de séparer les circuits intégrés élémentaires. Selon d'autres techniques, on procédera à un sciage par meule diamantée, ou encore à une découpe au laser. Chaque éclat de silicium qui représente ensuite un CI élémentaire porte, en raison de la petite taille des premiers CI, le nom de *puce*.

Fig. 3. — Aspect du MOS terminé, vu en coupe, avec sa protection par oxyde.



Bien entendu, des vérifications et mesures suivent toutes ces opérations, en particulier les mesures sous pointes : des pointes extrêmement fines viennent palper les CI sur le wafer, automatiquement, CI après CI. Les circuits reconnus défectueux sont marqués, par un jet d'encre par exemple.

Les puces des CI élémentaires sont ensuite encapsulées dans des boîtiers en métal, en céramique ou en plastique. En dernière étape sont effectuées des mesures finales, généralement sous la conduite d'un ordinateur qui dirige le programme des essais et procède aux tris et aux classements éventuels.

On verra par la suite que cette série d'opérations connaît bien des variantes, en fonction des technologies adoptées. On peut constater, par contre, que cette fabrication est plus simple que celle des circuits intégrés bipolaires, comme le résume le tableau B.

2. L'IMPLANTATION IONIQUE

Le dopage des semiconducteurs s'effectue généralement par diffusion gazeuse dans un four à haute température (1 100 °C environ). Or, cette technique ne permet pas une maîtrise parfaite de tous les paramètres du dopage, aussi a-t-on recherché d'autres méthodes. Ainsi, l'implantation ionique, qui tend à se développer depuis les années 70, semble devoir supplanter à terme, la diffusion gazeuse. Cependant, elle restait encore employée conjointement avec celle-ci jusqu'en 1978

L'implantation ionique consiste à bombarder le silicium à doper avec des ions du dopant. Accélérés par une tension précise qui leur communique l'énergie nécessaire — de quelques kilo-électronvolts à quelques centaines de kilo-électronvolts — ces ions pénètrent dans le cristal. Après dopage, le silicium est recuit à température relativement basse (500 à 800 °C) afin de restructurer le réseau cristallin.

Les avantages du procédé résultent du fait que :

- l'énergie des ions est parfaitement contrôlée;
- la dose d'ions par unité de surface également (sa mesure est une mesure d'intensité de courant);
- la température de recuit affecte bien moins le circuit car elle est notablement plus basse que la température dans un four de diffusion gazeuse (600 °C contre 1 100 °C). Elle interviendra donc bien moins sur les traitements déjà assurés;
- le dopage est très directif. Le diamètre du faisceau est compris entre 2 et 50 nm;
- il est également parfaitement reproductible;
- le dopage peut être assuré même à travers une couche isolante (silice, par exemple).

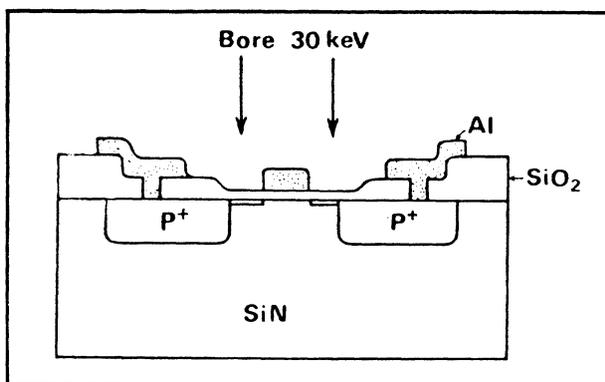


Fig. 4. — Implantation des drain et source au bore, sous 30 keV selon la technique auto-alignée proposée par Hughes.

Dernière particularité à porter à l'avantage de cette technique : elle permet le dopage de matériaux dans lesquels les diffusions sont quasi impossibles (le diamant par exemple).

L'implantation ionique permet la fabrication de très nombreux types de composants : transistors bipolaires, diodes, diodes à capacité variable, détecteurs de particules, photodiodes à avalanche, résistances de valeur élevée pour CI, MOS...

Les circuits intégrés MOS ainsi produits portent souvent le nom de I/MOS, ou IMOS pour : MOS implantés et l'implantation ionique est souvent désignée par le sigle I^2 qui rappelle ses initiales.

Les limitations en fréquence des MOS traditionnels viennent essentiellement des imperfections des techniques de masquage et de diffusion avec lesquelles il est difficile d'éviter les chevauchements du drain et de la source sur le gate. Le MOS implanté ne présente pas ces inconvénients, aussi les capacités interélectrodes sont nettement réduites, faisant apparaître un gain d'au moins deux sur les temps de commutation.

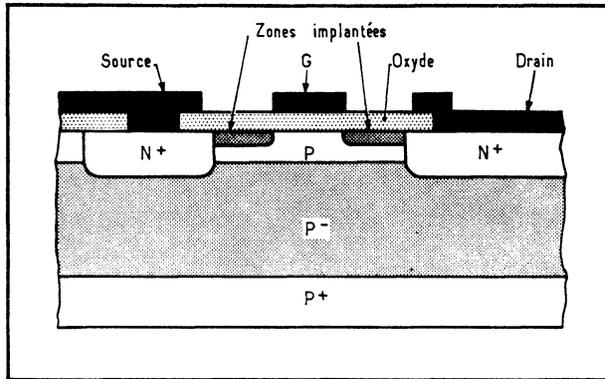
D'autre part, la tension de seuil du MOS est déterminée, en particulier, par l'épaisseur et la constante diélectrique de l'isolant de gate, et la densité de charge par unité de surface dans la zone de déplétion. Ce seuil peut alors être abaissé de manière à assurer une compatibilité avec les systèmes TTL; sa valeur typique est de 1 à 2 V.

Enfin, dernier avantage, les techniques d'implantation permettent de réaliser des MOS à appauvrissement et à enrichissement sur le même substrat. Leur combinaison détermine des portes dont le produit puissance consommée par rapidité est deux fois plus grand que celui des réalisations diffusées.

Technique d'implantation des drain et source

L'une des premières applications de l'implantation ionique à la technologie MOS semble due à l'équipe de R. BOWER, à la société *Hughes*. Elle était présentée à l'*International Electron Devices Meeting*, à Washington, en 1966, et visait une réalisation « auto-alignée » (fig. 4).

Fig. 5. — Technique *Philips* d'implantation ionique. A nouveau, on retrouve une structure auto-alignée. Le canal est situé dans la couche implantée, tandis que la partie principale de la région d'appauvrissement du drain est placée dans un matériau P- à forte résistivité obtenu par croissance épitaxiale sur un substrat P+, ce qui donne une faible capacité de drain.



L'électrode métallique de gate, plus étroite que l'espace drain-source, sert de masque au bombardement d'ions bore qui implantent la source et le drain à travers la couche mince d'oxyde. Ainsi, ces zones arrivent à l'aplomb du gate, sans recouvrement comme avec la diffusion gazeuse. Les zones métalliques ou d'oxyde épais arrêtent, elles, les ions accélérés. On aboutit ainsi à un auto-alignement des électrodes qui n'exige aucune opération supplémentaire de gravure, ni aucune modification du processus de fabrication. La dernière opération consiste en un recuit final, assuré à environ 500 °C. Le gain en rapidité du circuit est de l'ordre de 2 à 4 par rapport aux éléments classiques.

L'implantation peut être effectuée avec un accélérateur équipé d'une source haute fréquence dans laquelle est ionisé du trichlorure de bore. Après un tri magnétique, le faisceau d'ions bore est soumis à un balayage électrostatique XY et permet d'obtenir une zone implantée homogène sur la pastille de silicium qui reste à température ambiante.

La méthode *Philips* est très semblable (fig. 5). Là aussi, les ions sont arrêtés par l'aluminium du gate; ce sont des ions bore ou phosphore selon que l'on veut créer un MOS à canal P ou à canal N. Le canal peut être réduit à 3 μm de longueur ce qui a permis de fabriquer des MOS oscillant à 800 MHz.

Implantation de résistances

Selon ce processus, les résistances de couche des zones implantées sont assez élevées et atteignent 1 à 2 $k\Omega/\square$. On peut cependant, de la même manière, aboutir à des résistances de fortes valeurs, comprises entre 10 et 50 $k\Omega/\square$, qui peuvent servir de résistances de charge.

L'implantation du canal

Au lieu d'implanter les drain et source, on peut implanter le canal, toujours à travers la mince couche d'oxyde de gate. Le processus développé ainsi par *Sprague-*

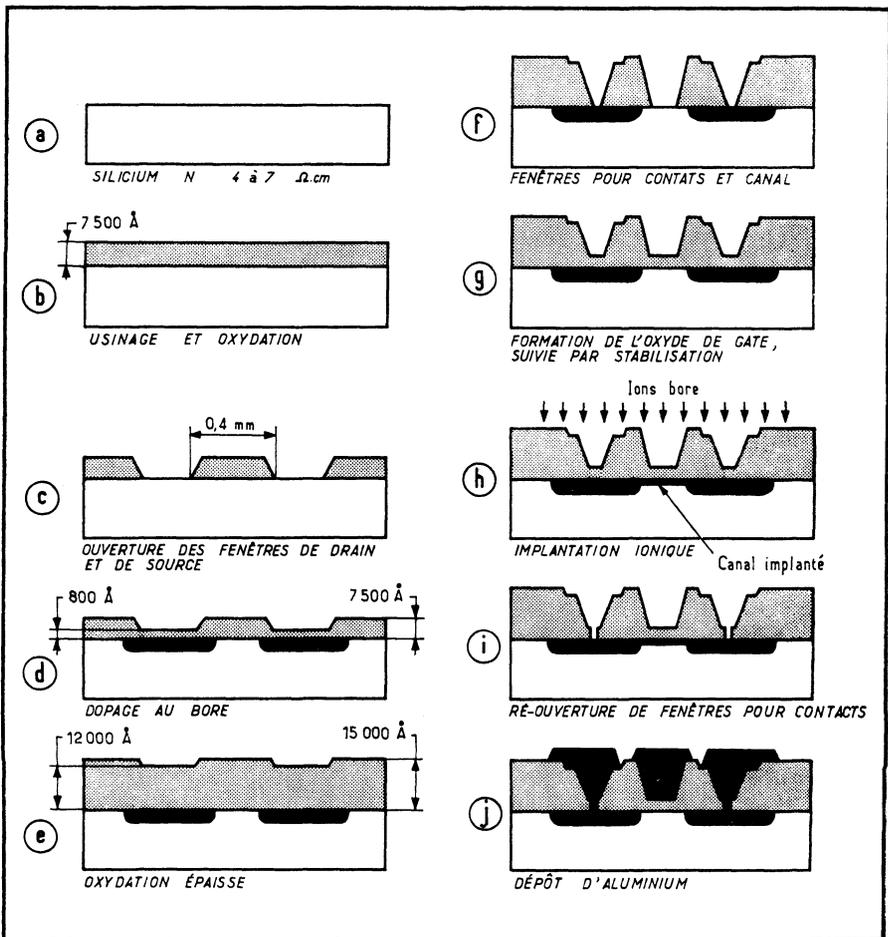


Fig. 6. — Film de la fabrication d'un MOS par implantation ionique, selon la technologie mise au point par *Sprague-Mostek*.

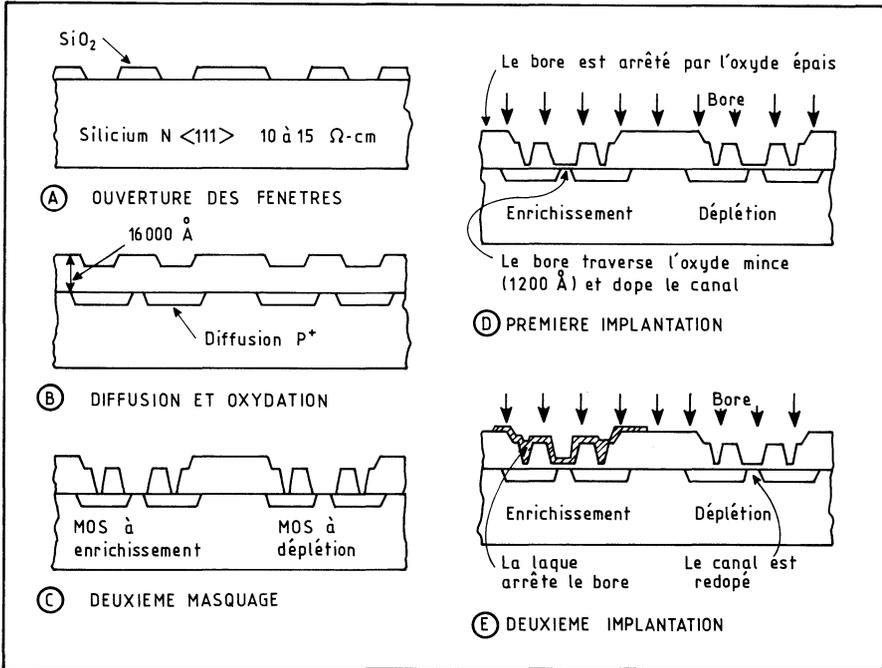


Fig. 7. — Fabrication d'un inverseur avec charge à déplétion.

Mostek est décrit par la figure 6; en France, le *LETI*, en particulier, recourt également à cette technique, auto-alignée elle aussi, mais qui permet de réaliser aussi bien des MOS à enrichissement que des MOS à déplétion (les opérations n'étant toutefois pas simultanées pour l'ensemble des deux types, mais séparées).

En effet, on n'a vu que la tension de seuil V_{TH} était donnée par :

$$V_{\text{TH}} = \Phi_{\text{MS}} - \frac{Q_{\text{SS}}}{C_{\text{O}}} - \frac{Q_{\text{B}}}{C_{\text{O}}} + 2\Phi_{\text{F}}$$

où Q_{B} représente les charges dans la zone de déplétion du substrat, située juste au-dessous de la surface d'inversion du canal. Or, l'implantation du canal permet de modifier Q_{B} d'une manière continue et dans une gamme étendue : on peut alors ajuster la tension de seuil du MOS à une valeur précise prédéterminée.

Bien plus : si l'on introduit des atomes de type accepteur (bore), on déplace V_{TH} vers les tensions positives. A l'inverse et avec l'implantation d'atomes donneurs dans le substrat, tel le phosphore, on déplace V_{TH} vers les valeurs négatives. Ainsi, on peut agir à volonté sur la tension de seuil des MOS, aussi bien à canal N qu'à canal P.

Par contre, la tension d'inversion V_{TH} qui est la tension de seuil déterminée par l'oxyde épais et jouant sur les MOS parasites, n'est pas modifiée. Si l'on voulait agir sur le rapport $V_{\text{TF}}/V_{\text{TH}}$, on pourrait d'ailleurs modifier la concentration en porteurs de substrat en surface à l'aide d'un « prédépôt » d'ions implantés hors des zones utiles de source, gate et drain.

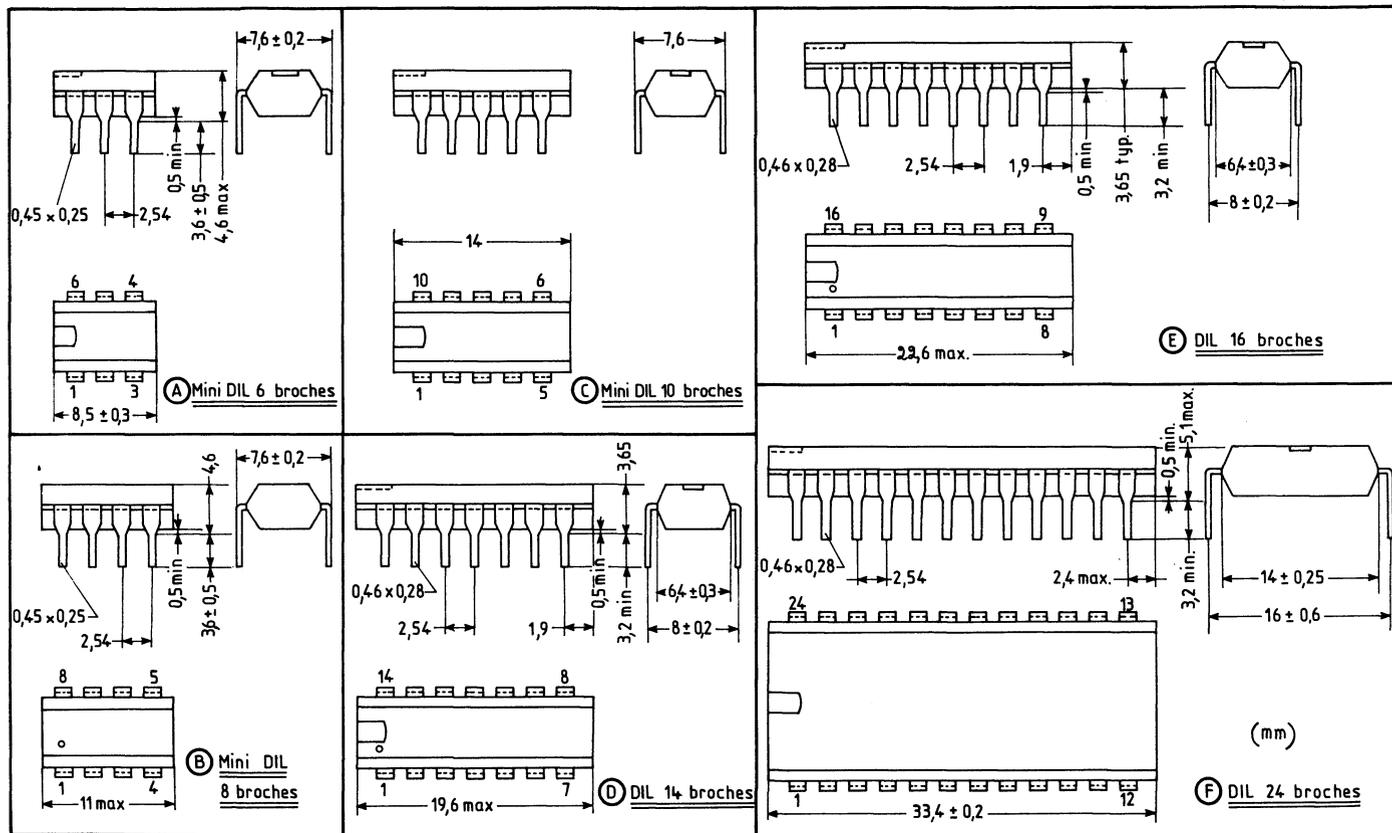


Fig. 8. — Quelques types de boîtiers « dual-in-line » (DIL) en plastique.

Ainsi, une concentration de porteurs de $10^{15}/\text{cm}^3$ à $10^{16}/\text{cm}^3$ accroît d'environ 15 V la tension de seuil parasite V_{TF} pour une épaisseur d'oxyde de 1 700 nm.

Fabrication d'inverseurs avec charge à déplétion

La réalisation d'un inverseur associant un MOS canal P à enrichissement à un MOS, canal P également, mais à déplétion, obéit aux règles qui viennent d'être énoncées mais exige un masquage supplémentaire (fig. 7).

La première implantation P, au bore, vise les deux MOS simultanément. Puis le MOS à enrichissement est masqué afin que la seconde implantation de bore n'atteigne que le MOS à déplétion.

Fabrication de caissons pour MOS

Avec les MOS complémentaires, ou CMOS, on a besoin de caissons de type P pour le MOS à canal N. Un tel caisson peut être réalisé sous forme de prédépôt, par implantation ionique donc.

L'usinage ionique

La gravure ionique, qui est un cas particulier de la pulvérisation, repose sur le transfert de la quantité de mouvement des ions incidents aux atomes de la surface bombardée. C'est donc là une autre application du bombardement ionique, venant remplacer ou compléter la gravure chimique.

3. LA PRÉSENTATION DES CI

Les puces de silicium peuvent être encapsulées dans des boîtiers, ou encore présentées sous forme de puces nues selon des technologies particulières.

Boîtiers

On utilise des boîtiers :

- pour l'essentiel, en plastique, type à double rangée de connexions, désormais fiables et économiques (fig. 8);
- en céramique pour les besoins de haute fiabilité (fig. 9);
- métalliques pour assurer l'étanchéité.

Ces boîtiers sont du type :

- surtout à *double rangée de connexions* (« dual-in-line package », soit *DIL* ou *DIP*). Les *DIL* existent jusque vers 40 broches;
- *plats (flat pack)*, destinés surtout aux applications militaires (fig. 10);
- du type *transistor*, peu utilisé désormais en numérique.

Les puces nues

L'encapsulation coûtant jusqu'à la moitié du prix du CI (dans les cas extrêmes!), on a songé à utiliser telles les puces nues. Pour cela, ces puces peuvent être :

- montées sur des supports intermédiaires *LID*, *Cératab*, ou autres;
- présentées en *beam-leads*, ou puces à conducteurs-poutres (fig. 11);
- présentées en *flip-chips*, ou *pastilles à protubérances* (fig. 12);
- présentées en *spider-bonding*, ou en *pattes d'araignée*, comme pour le montage en DIL plastique (fig. 13).

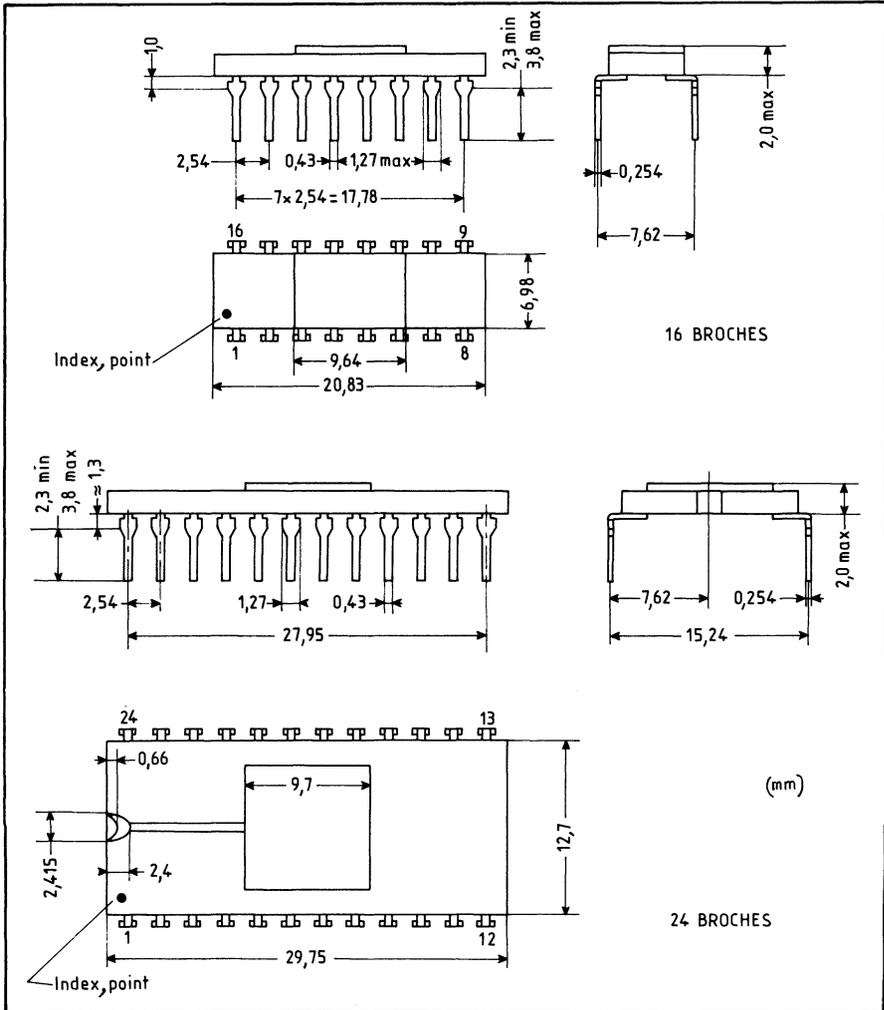


Fig. 9. — Quelques boîtiers DIL en céramique.

4. CIRCUITS SUR FILM

Parce que manuelles, les opérations d'encapsulation sont très onéreuses; elles ont ainsi justifié l'implantation d'usines dans des pays à bas prix de main-d'œuvre. Or, il existe plusieurs méthodes permettant d'automatiser cette étape de la production, pourvu que les quantités soient suffisantes. L'une de celles-ci, peut-être la première annoncée, est la technique *Minimod*, élaborée par *General Electric* en 1971 et reprise par *Texas* ultérieurement. Dès 1975, l'équipe parisienne de *Honeywell-Bull* annonçait même une chaîne totalement opérationnelle et permettant d'automatiser le traitement des puces, depuis leur réalisation sur la tranche de silicium jusqu'à leur montage sur circuit hybride via un support-film.

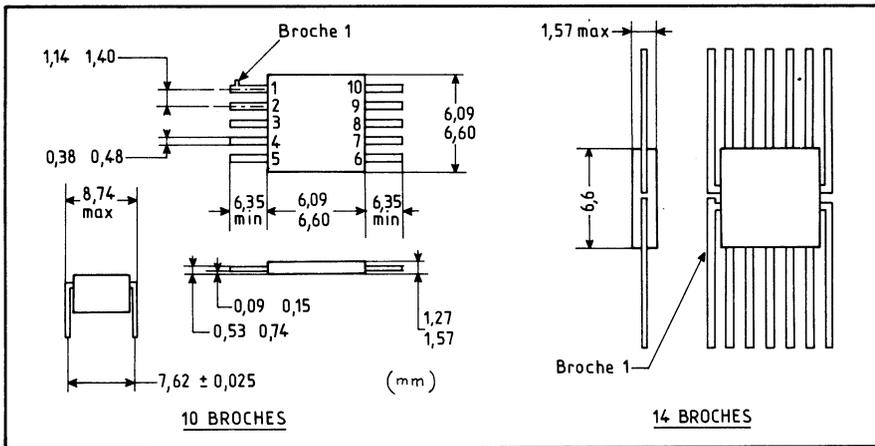


Fig. 10. — Exemples de boîtiers plats.

Fig. 11. — Montage d'un beam-lead sur un circuit hybride. Les « poutres », partie intégrante de la puce semi-conductrice, assurent sa fixation mécanique et sa connexion électrique.

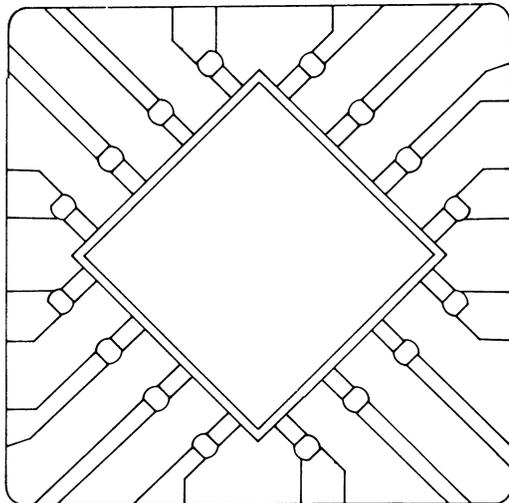


Tableau C. — Quelques techniques d'encapsulation

	Technique	Pastille nue	Beam-lead	Flip-chip	Spider-bonding
FIABILITÉ	Herméticité	Non	Oui	Non	Non
	Protection de la surface	Moyenne	Excellente	Moyenne	Moyenne
	Fiabilité des soudures	Faible	Excellente	Moyenne	Moyenne
	Possibilité d'inspection des soudures	Oui	Oui	Non	Oui
	Procédé de fabrication	Standard	Standard jusqu'à diffusion émetteur	Standard jusqu'à métallisation	Standard sauf pour empla- cement soudure des fils de sorties
	Caractéristiques thermiques	Excellentes	Excellentes	Moyennes	Excellentes
	Fiabilité totale	Faible	Moyenne	Bonne	Moyenne
COÛT	Coût de la pastille	Peu élevé	Élevé	Élevé	Moyen
	Possibilité de réparation	Oui	Oui	Oui	Non
	Possibilité de réaliser du multipastille dans un seul boîtier	Très faible	Moyenne ou excellente	Excellente	Faible
	Coût au niveau du système	Très élevé	Moyen ou faible	Faible	Élevé

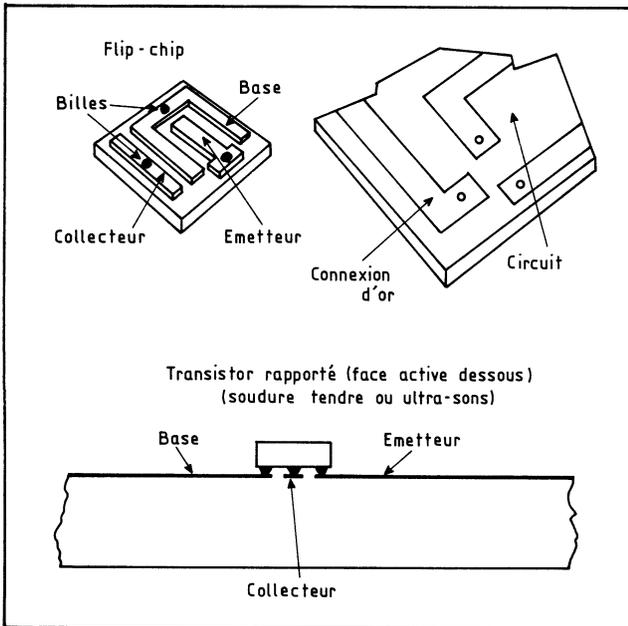


Fig. 12. — Le transistor à protubérances, ou flip-chip, est monté sur l'hybride, face active contre le substrat (en haut); vue en coupe de la puce soudée (en bas).

Ce mode de présentation des circuits intégrés sur film s'est fortement développé, sous de multiples variantes, en raison précisément de l'automatisation qu'il permet : aussi bien pour le test que pour le report sur circuit hybride ou montage dans un boîtier.

5. TECHNIQUES ÉVOLUÉES DE MASQUAGE

La recherche de circuits toujours plus complexes et plus denses a mis en évidence les limites de certaines technologies. Ainsi, le masquage et l'insolément,

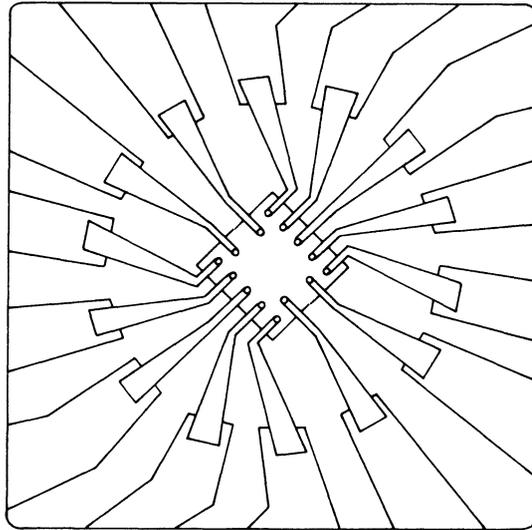


Fig. 13. — Les connexions « en araignée » du spider-bonding. La puce semiconductrice est montée, face active en l'air; elle doit être fixée sur le substrat hybride avant soudure de ses connexions.

même en ultraviolet, ne peuvent fournir qu'une résolution de l'ordre de 1 à 2 μm au mieux; leur processus d'emploi est le suivant :

- on réalise un maître-masque par procédé photographique; il y a autant de maîtres-masques que de nombre de masquages;
- par contact ou par duplication, on reproduit ce maître-masque en masques de travail, en chrome sur verre généralement;
- chaque masque est reporté sur les tranches de silicium oxydées et recouvertes d'une résine photosensible (avec une précision de superposition de l'ordre du micron), toujours donc par voie photographique.

La longueur d'onde des photons limite la définition; avec les électrons utilisés en bombardement sous vide, la longueur d'onde comme les phénomènes de diffraction permettent une amélioration de la définition d'un ordre de grandeur, comme en témoigne le générateur de masques de *Thomson-CSF*, qui préfigure peut-être le type d'appareil de production auquel la microélectronique se devra de recourir.

Le masqueur électronique est utilisé pour la production de masques de haute précision (0,1 μm) sur plaques photographiques ou sur plaque au chrome, ou directement sur semiconducteur.

Remplaçant un coordinatographe, un appareil de réduction, une machine de répétition par pas et une machine d'alignement, ce masqueur exécute en une seule opération l'ensemble du traitement.

Le principe de fonctionnement est le suivant. Un motif du masque est tracé en balayant le faisceau électronique sur l'échantillon recouvert de résine électrosensible, suivant un programme défini directement à partir de données numériques.

Pour couvrir toute la surface du masque, l'échantillon est déplacé sous le faisceau par des moteurs pas-à-pas. La position de la platine porte-échantillon est mesurée à l'aide de deux interféromètres à laser, et l'erreur de position ainsi relevée est corrigée par un léger déplacement du faisceau électronique.

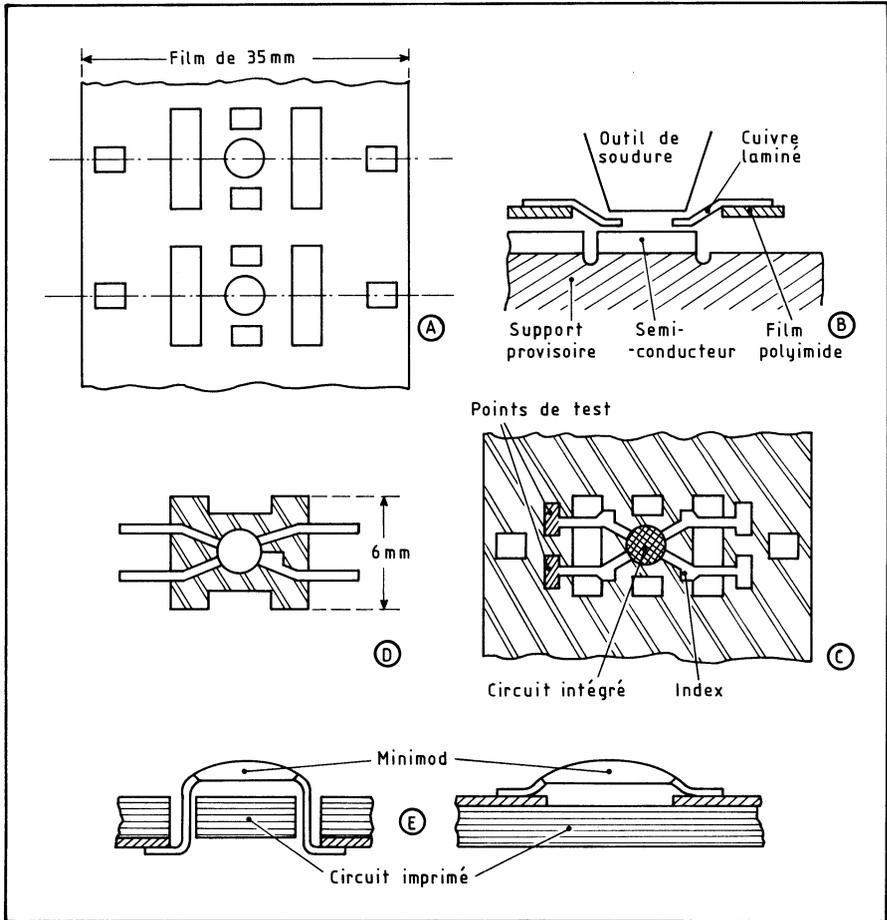


Fig. 14. — Montage d'un circuit intégré sur film perforé (a). En b, principe de la soudure « Multi-bond », mise au point par General Electric (« G.E. ») pour boîtiers miniMod. Quatre points de test apparaissent sur le film (deux à gauche, ici hachurés et deux à droite), et facilitent grandement les contrôles automatiques (en c). En d, un circuit intégré détaché du film-support. En e, enfin, montage direct d'un C.I. sur une plaque imprimée.

L'inconvénient du procédé, qui dérive du principe adopté avec le microscope électronique, réside dans la durée d'exposition qui est relativement longue. Il faut utiliser, bien entendu, une résine sensible aux électrons (et non plus aux photons).

Une autre méthode consiste en l'utilisation du « flying spot ».

La reproduction sur film photographique à l'échelle est assurée, ici, à partir d'un maître-masque disposé entre un tube cathodique qui assure le balayage et un

photomultiplicateur. Le tube cathodique est synchronisé avec le canon à électrons; la lumière reçue par le photomultiplicateur à partir du tube dont le spot (flying spot) balaye tout l'écran est amplifiée et transmise à un dispositif à seuil qui commande ou non les bobines de blocage.

Lithographie aux rayons X

Une autre technique peut être exploitée : la lithographie aux rayons X. Elle autorise une résolution de 0,1 micron, une profondeur de champ de 10 microns; elle est insensible à la présence de particules de poussières ou de contaminants; enfin, elle peut être utilisée avec des résists positifs ou négatifs.

Dans la lithographie à rayons X (ou « radiolithographie »), un masque est placé au-dessus d'un substrat revêtu d'un film polymère (résist) radio-sensible. Le masque contient des zones qui absorbent fortement les rayons X mous. Une source de rayons X, excitée par un faisceau d'électrons, projette l'ombre de ces zones absorbantes du masque sur le film polymère.

Masquage au laser

La reproduction des masques gagne en vitesse si l'on recourt au laser. Par exemple, aux laboratoires *Philips* de Hambourg, un faisceau optique, émis par un laser à argon de 1 watt, traverse un modulateur électro-optique, puis un défecteur également électro-optique, avant d'être envoyé sur la plaque à graver. Celle-ci est recouverte d'une fine couche métallique de chrome sur 1 000 angströms, que le faisceau laser évapore localement, en y laissant un trou microscopique. Le diamètre du faisceau, à l'impact, est de 1 à 2 micron; cette technique permet de réaliser près d'un million de points par millimètre carré, en deux secondes.

Les chercheurs de *Bell Laboratories* utilisent une autre technologie : le faisceau laser trace le masque directement sur un film photo-sensible; le modulateur est ici acousto-optique, tandis que le défecteur est mécanique (il est constitué par un rotor à forme décagonale, chacun des côtés de ce rotor supportant un miroir). Un masque complet de circuit intégré est alors produit en 12 minutes.

Aux *Western Electric Co.* (Princeton), un masqueur à laser a été mis au point très récemment. Il utilise des défecteurs galvanométriques, et peut soit travailler sous la commande d'un ordinateur (PDP-15) pour créer de nouveaux masques, soit effectuer simplement un balayage pour effectuer la réplique d'un maître-masque.

CHAPITRE V

L'ARSENAL DES TECHNOLOGIES

1. INTRODUCTION

Jusqu'en 1975, la logique bipolaire, avec à sa tête la TTL, a dominé le marché des circuits intégrés. C'est la raison pour laquelle on a cherché à rendre compatibles avec la TTL toutes les familles ultérieures, et essentiellement celles à MOS. En effet, des familles compatibles permettent :

1. Soit d'associer MOS et bipolaires dans un même montage sans rendre trop complexes les problèmes d'alignement des niveaux et tensions entre ces types de circuits.
2. Soit de transposer plus aisément dans de nouvelles familles tout ou partie de montages existants ou nouveaux.
3. Soit, enfin, de mélanger MOS et bipolaires dans un même et unique CI.

Or, comme le montre le tableau A, les caractéristiques en tension et en niveaux logiques notamment des premiers MOS étaient fort éloignées de celles de la TTL.

Tout mariage MOS-TTL, quasi-obligatoire au début de l'utilisation des MOS, imposait donc le recours à des interfaces chargées de transposer les niveaux; celles-ci compliquaient les montages et intervenaient également dans leur prix.

La solution a alors consisté à rechercher des familles MOS capables de s'aligner sur les caractéristiques de la TTL, et pour ce faire, à fabriquer des MOS à faible tension de seuil.

En effet, le MOS fondamental à enrichissement, canal P, a une tension de seuil V_{TH} de l'ordre de -4 V. Si l'on réussit à la réduire ce V_{TH} , on réduira du même coup les tensions d'alimentation et les niveaux logiques qui se rapprocheront ou rejoindront ceux de la TTL.

Tableau A. —
Le MOS fondamental comparé au bipolaire (TTL)

CARACTÉRISTIQUES	TTL 54/74	MOS Canal P ENRICHISSEMENT
Tensions d'alimentation		
V_{CC} (collecteur)	+ 5 V	
V_{DD} (drain)		-13V
V_{EG} (gate)		-27V
Niveaux logiques		
0	0V	0 V
1	> 3,5V	-12V
Gammes de températures de service	-55 à +125°C	-20 à +85°C
Fréquences de travail	30 MHz	1 MHz

C'est pourquoi l'on parle de MOS à *haut seuil*, en traitant du MOS fondamental, et de MOS à *bas seuil* pour les familles MOS dont le V_{TH} est moindre et varie autour de 2 V. La première formule retenue a consisté à utiliser du silicium orienté $\langle 100 \rangle$ pour réduire cette tension de seuil.

Au préalable, rappelons qu'en bipolaire (et en MOS complémentaires également), on utilise la logique *positive* : la tension la plus élevée correspond au 1 logique, la tension la plus basse au 0. A l'inverse, avec les MOS monocanaux, c'est la logique *négative* qui intervient : le 1 logique est la tension la plus basse, le 0 logique la tension la plus élevée. On verra, au besoin, comment passer de la logique positive à la négative, ce qui se révèle d'ailleurs très simple.

2. SILICIUM $\langle 111 \rangle$

La structure cubique du silicium se prête à un découpage en tranches selon divers plans d'orientation, comme on l'a montré dans le chapitre consacré à la

fabrication du MOS fondamental. Celui-ci est produit à partir de silicium $\langle 111 \rangle$; la tension de seuil obtenue est environ de :

$$V_{TH} = -4 \text{ V (de } -3 \text{ à } -5 \text{ V)}$$

La tension d'inversion V_{TF} qui commande le MOS parasite est :

$$V_{TH} = -30 \text{ V à } -40 \text{ V}$$

Les tensions d'alimentation drain (V_{DD}) et gate (V_{GG}) répondent aux approximations pratiques suivantes :

$$V_{DD} \approx 3 V_{TH} \approx -12 \text{ V}$$

$$V_{GG} \approx 6 V_{TH} \approx -24 \text{ V}$$

3. SILICIUM $\langle 100 \rangle$

Si l'on utilise du silicium orienté $\langle 100 \rangle$, la charge en surface Q_{SS} qui intervient dans le calcul de V_{TH} est plus faible, et de ce fait, V_{TH} est réduit. On obtient pratiquement (avec $Q_{SS} < 1 \cdot 10^{11}$ charges/cm²) :

$$V_{TH} = -1,7 \text{ à } -2,2 \text{ V}$$

$$V_{TF} = -15 \text{ à } -20 \text{ V}$$

$$V_{DD} = 3 V_{TH} = -5 \text{ à } -7 \text{ V}$$

$$V_{GG} = 6 V_{TH} = -6 \text{ à } -12 \text{ V}$$

La première conclusion, c'est qu'on peut désormais adopter une valeur $V_{DD} = -5 \text{ V}$, ou mieux, porter le niveau de référence (masse) à $+5 \text{ V}$, et le V_{DD} à 0 V ; le V_{GG} sera alors de -12 V , par exemple. Le $+5 \text{ V}$ est désormais une tension commune aux MOS et à la TTL; le 12 V , lui, pourrait convenir aux amplificateurs opérationnels également, si besoin était... On peut même accepter à l'entrée des MOS les niveaux logiques caractéristiques de la TTL, à la limite toutefois.

Un autre avantage consiste en une réduction de la consommation. Il est normal, en effet, que si l'on réduit les tensions d'alimentation et si toutes autres choses restent égales, on aboutisse à une diminution de la puissance dissipée.

Par contre, les inconvénients du procédé sont de deux ordres :

1. La tension d'inversion sur l'oxyde épais est réduite, ce qui réduit en *valeur absolue* la marge de bruit sur les tensions d'alimentation. L'écart $V_{TF} - V_{GG}$ est en effet, dans le cas le plus défavorable, de :

- Pour le silicium $\langle 100 \rangle$:

$$20 \text{ V} - (12 + 5 \text{ V}) = 3 \text{ V}$$

- Pour le silicium $\langle 111 \rangle$:

$$30 \text{ V} - 27 \text{ V} = 3 \text{ V}$$

On notera qu'en valeur relative, l'avantage irait plutôt au silicium $\langle 100 \rangle$.

2. Mais il y a plus grave : la mobilité des porteurs est réduite; il s'ensuit que la vitesse de travail est moindre.

Comme, de plus, le procédé $\langle 100 \rangle$ exige une surface de silicium supérieure de $+20 \%$ environ, à performances égales, il n'a pas été largement développé. La fabrication des MOS $\langle 100 \rangle$ reste rigoureusement identique à celle des $\langle 111 \rangle$, au plan de coupe près.

4. UTILISATION DU NITRURE DE SILICIUM

Le nitrure de silicium bénéficie d'une constante diélectrique supérieure de près du double de celle du dioxyde de silicium : 7,5 au lieu de 3,9 environ. Il supporte donc des champs électriques plus élevés à épaisseur identique, c'est-à-dire qu'il résistera à la même tension sous une épaisseur moindre.

Or, si l'on réduit l'épaisseur de l'isolant qui sépare le gate du canal, on réduit du même coup la tension de seuil V_{TH} qui commande le MOS : de 4 à 5 V, elle peut passer à 2 V environ.

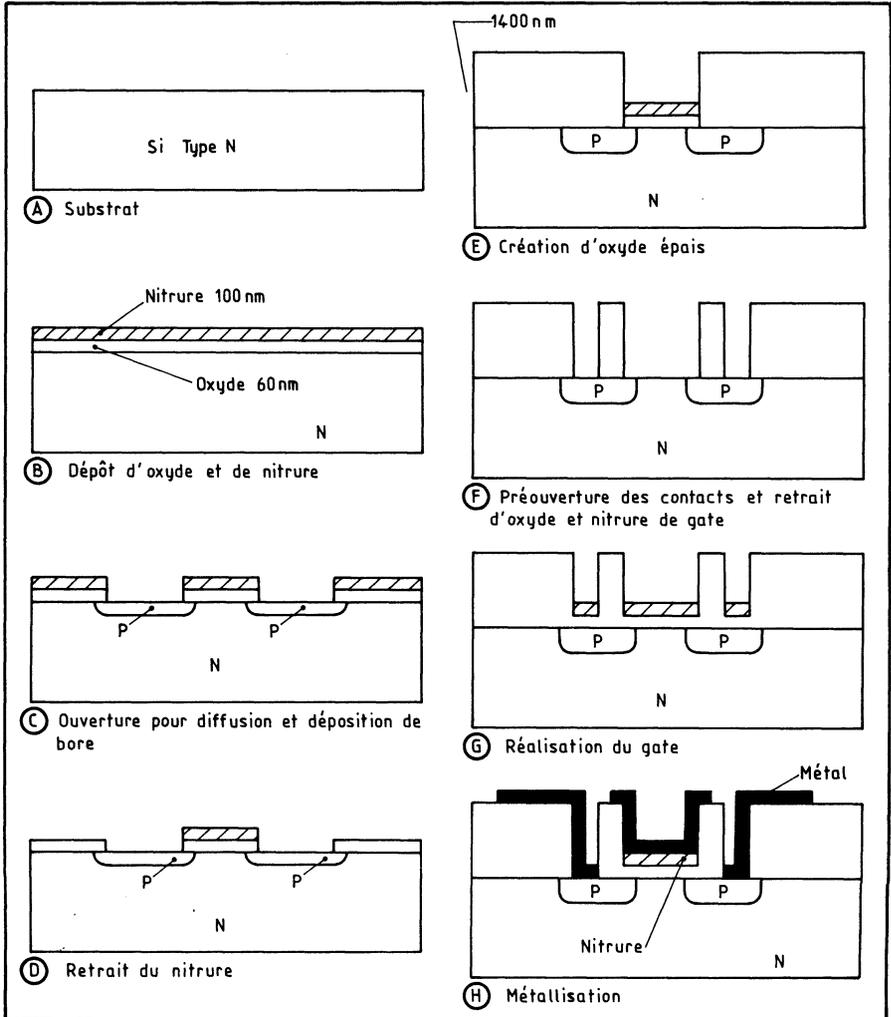


Fig. 1. — Fabrication du MOS au nitrure de silicium : l'autoalignement.

Il s'ensuit que les tensions d'alimentation peuvent être réduites, elles aussi : d'une excursion totale de plus de 27 V avec le MOS fondamental, on passe à 17 V en adoptant des tensions de +5 V et -12 V pour le MOS au nitrure de silicium. On remarquera que ces dernières sont précisément les tensions utilisées avec les CI bipolaires (le 5 V pour la DTL/TTL, le 12 V pour les comparateurs, les amplificateurs opérationnels...). Accessoirement, la puissance consommée $P = VI$, elle aussi est réduite.

D'autre part, le nitrure s'oppose à la migration des ions sodium jusqu'à environ 200 °C, ce qui contribue à améliorer la stabilité à long terme des dispositifs, en tension de seuil et en courants de fuite. C'est la raison pour laquelle on peut faire travailler les nouveaux dispositifs MOS jusqu'à 125 °C sans risque de dérive. Pratiquement, cela se traduit par le fait qu'avec les éléments au nitrure, la gamme de travail devient -55 à +125 °C.

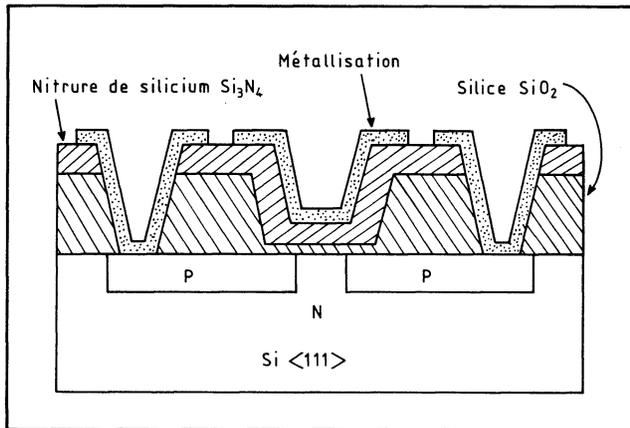


Fig. 2. — MOS au nitrure sur silicium 111.

Enfin, on peut également rendre compatibles les niveaux logiques des MOS au nitrure avec ceux des logiques DTL/TTL. Cet avantage se traduit par la suppression des interfaces qui, autrement, étaient nécessaires pour relier ces familles différentes.

Pour fabriquer un CI/MOS au nitrure, on réalise un sandwich de nitrure et de dioxyde de silicium, comme le montre la figure 1. Malheureusement, en effet, le nitrure déposé directement sur le wafer en silicium ferait apparaître des charges indésirables près de l'interface. De ce fait, la constante diélectrique moyenne est pratiquement de l'ordre de 5,6.

Un tel dispositif voit sa résistance de canal R_{on} réduite par rapport à celle du MOS conventionnel et, puisque les constantes de temps sont réduites, les fréquences de travail s'accroissent. On obtient les caractéristiques générales suivantes :

$$V_{TH} = -2 \text{ à } -2,5 \text{ V}$$

$$V_{TF} = -30 \text{ à } -50 \text{ V}$$

$$V_{DD} = -6 \text{ à } -7,5 \text{ V}$$

$$V_{GG} = -12 \text{ à } -15 \text{ V}$$

Cette structure, qui permet de revenir au silicium orienté < 111 > (fig. 2), exige plus d'étapes de fabrication que le < 111 > fondamental. Elle autorise un interfaçage direct avec la TTL, car l'excursion totale de tension peut être distribuée entre un $V_{DD} = +5 \text{ V}$ et $V_{GG} = -12 \text{ V}$ comme avec le procédé < 100 >. En

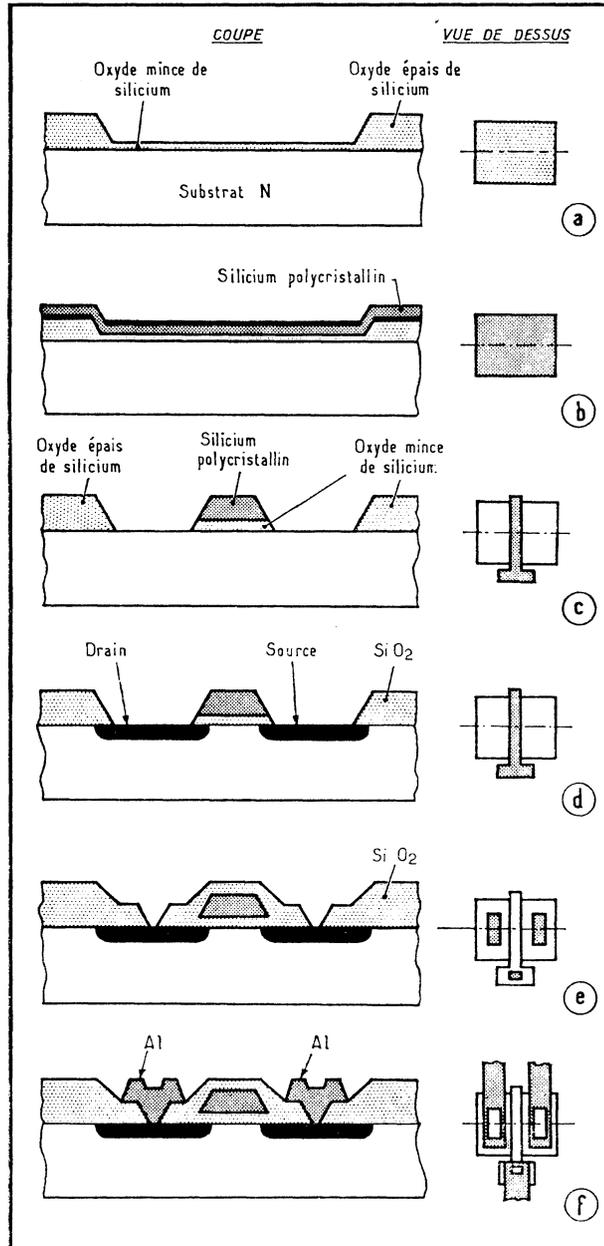


Fig. 3. — La fabrication d'un MOS à gate au silicium (processus *Fairchild*). Le silicium polycristallin du gate est fortement dopé; sa connexion de sortie apparaît dans la série de vues de dessus.

raison de la protection supplémentaire qu'apporte le nitrure, la fiabilité est améliorée; en outre, la densité d'intégration s'accroît car la dimension des éléments peut être réduite.

Les premiers CI au nitrure ont été commercialisés dès 1969 par *General Instrument* sous la dénomination de MTNS (*metal, thick oxide, nitride, silicon*), puis ensuite de *GIANT* (*General Instrument Advanced Nitride Technology*).

5. GATE AU SILICIUM

L'épaisseur du diélectrique de gate n'est pas le seul facteur qui intervient sur la tension de seuil : il faut tenir compte, en second lieu, de la nature du gate.

En effet, la différence entre le travail de sortie (pour extraire un électron) du gate et celui du substrat constitue un potentiel de contact Φ_{MS} dont dépend V_{TH} , comme on l'a vu précédemment. La formule donnant V_{TH} est, rappelons-le :

$$V_{TH} = -\frac{Q_{SS}}{C_O} - \frac{Q_B}{C_O} + \Phi_{MS} + 2\Phi_F$$

On constate ainsi que, pour réduire V_{TH} , il faut réduire le potentiel de contact Φ_{MS} , qui est de l'ordre de $-0,35$ V avec la structure aluminium-silicium pour une résistivité de silicium comprise entre 1 et $10 \Omega/\text{cm}$, type P. Or, on peut remplacer l'aluminium *conducteur*, par du silicium polycristallin fortement dopé : ce faisant, le potentiel de contact passe à $+0,6$ V dans les mêmes conditions, soit une différence de l'ordre du volt.

Avec une structure à gate d'aluminium, on avait calculé :

$$V_{TH} = -1,05 - 1,4 - 0,35 - 0,5 = -3,3 \text{ V}$$

Si l'on passe au gate au silicium, on trouvera dans les mêmes conditions :

$$V_{TH} = -1,05 - 1,4 + 0,6 - 0,5 = -2,35 \text{ V}$$

En dopant diversement ce silicium, on agit d'ailleurs sur la tension de seuil qui peut être ajustée à volonté.

Les premiers transistors MOS à gate au silicium (les *silicon gate MOS* des Américains) sont dus aux travaux, de J. C. SARACE et de son équipe, aux *Bell Telephone Laboratories*.

Le processus de fabrication d'un MOS à gate en silicium est développé dans la figure 3, selon la méthode élaborée dès 1967 par *Fairchild*. On part d'un substrat N sur lequel on fait croître une couche épaisse d'oxyde : puis on ouvre des fenêtres, une par transistor à fabriquer, et l'on remplace l'oxyde épais par de l'oxyde mince : c'est là qu'intervient donc le premier masquage (en *a*).

Après quoi, toute la surface du wafer (du disque de silicium sur lequel sont réalisés les CI) est recouverte de silicium polycristallin (*b*). Un second masquage délimite la zone de gate, et par la même occasion, un éventuel premier plan d'interconnexions. L'oxyde mince et le silicium polycristallin sont éliminés, sauf à l'emplacement du gate, comme le montre la figure 3 *c* : cet emplacement circonscrit très précisément l'action de la diffusion du bore qui, sans masquage supplémentaire, est immédiatement assurée (*d*). Le silicium polycristallin du gate est donc, lui aussi, dopé à ce moment, et sa résistivité tombe à $1 \Omega/\text{cm}$ environ.

Une première différence importante apparaît donc à ce niveau avec les structures classiques : c'est ce que l'on a voulu préciser figure 4. On voit que le chevauchement du gate en aluminium sur le drain et la source a disparu avec la technique du gate en silicium, d'où il s'ensuit une réduction notable des capacités.

Mais revenons au processus technologique. Le bore, qui diffuse relativement vite dans le silicium, laisse pratiquement intact l'oxyde de silicium; cette diffusion (la seule dans tout le processus de fabrication) vise non seulement à créer le drain et la source, mais également à doper le gate pour l'amener à une faible résistivité, d'environ $1 \Omega/\text{cm}$.

Suivent une nouvelle oxydation et un masquage, le troisième, qui prépare le dégagement des prises de contact de drain et de source (*e*). Une couche d'aluminium est alors déposée sur toute la surface du wafer, un quatrième et dernier masquage intervenant pour n'en laisser subsister que ce qui est nécessaire au second plan d'interconnexions (*f*).

La technologie à gate au silicium s'est rapidement développée en raison de son intérêt. On notera d'ailleurs qu'elle est compatible avec d'autres processus technologiques, comme on l'évoquera ultérieurement. Notons à son actif :

1. *La notion de rendement en fabrication*, qui est essentielle en microélectronique, intervient en tout premier lieu : or, les rendements sont, ici, supérieurs à ceux obtenus avec la technologie fondamentale parce que l'oxyde du gate, point sensible de l'élément, est aussitôt recouvert par du silicium qui le protège de toute contamination ultérieure.

2. *On aura remarqué qu'avec ce processus, le « chevauchement » du gate sur les zones de source et de drain est, sinon totalement supprimé, du moins réduit au minimum* que permet la diffusion.

En effet, celle-ci est assurée après réalisation du gate : *on dit que ce processus est à auto-alignement des électrodes*. Le recouvrement peut alors être réduit jusqu'au micromètre, alors qu'il est normalement de 5 à 10 μm avec la technologie de base.

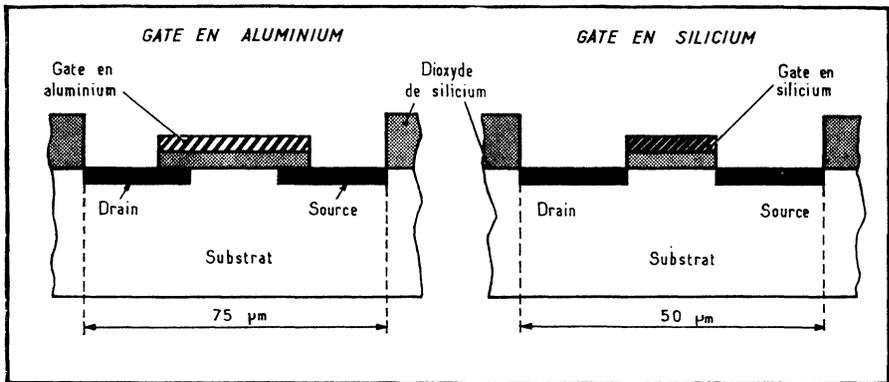


Fig. 4. — Structures comparées d'un MOS classique et d'un MOS à gate en silicium. Les électrodes de la structure CIS (« *conducteur-isolant-semiconducteur* ») sont auto-alignées ce qui évite tout chevauchement du gate sur le drain et la source, donc réduit considérablement les capacités de gate.

3. *Le gate peut, de plus, être plus court*. Ces deux raisons contribuent à une réduction des capacités du transistor, donc à un accroissement du gain et, ce qui est encore plus important, les vitesses de travail augmentent sensiblement.

4. *La tension de seuil est abaissée à environ 2 V*. On peut à nouveau réduire les tensions d'alimentation et les rendre, ainsi que les niveaux logiques, compatibles avec les logiques bipolaires DTL/TTL.

5. *On a vu que le procédé de fabrication fournit deux couches d'interconnexions* : c'est là une qualité des plus appréciables en LSI. Si l'on inclut la couche d'interconnexions possible due au substrat, on aboutit à un total de 3 couches.

Les premiers CI au nitrure ont été commercialisés dès 1969 par *General Instrument* sous la dénomination de MTNS (*metal, thick oxide, nitride, silicon*), puis ensuite de *GIANT* (*General Instrument Advanced Nitride Technology*).

5. GATE AU SILICIUM

L'épaisseur du diélectrique de gate n'est pas le seul facteur qui intervient sur la tension de seuil : il faut tenir compte, en second lieu, de la nature du gate.

En effet, la différence entre le travail de sortie (pour extraire un électron) du gate et celui du substrat constitue un potentiel de contact Φ_{MS} dont dépend V_{TH} , comme on l'a vu précédemment. La formule donnant V_{TH} est, rappelons-le :

$$V_{TH} = -\frac{Q_{SS}}{C_O} - \frac{Q_B}{C_O} + \Phi_{MS} + 2\Phi_F$$

On constate ainsi que, pour réduire V_{TH} , il faut réduire le potentiel de contact Φ_{MS} , qui est de l'ordre de $-0,35$ V avec la structure aluminium-silicium pour une résistivité de silicium comprise entre 1 et $10 \Omega/\text{cm}$, type P. Or, on peut remplacer l'aluminium *conducteur*, par du silicium polycristallin fortement dopé : ce faisant, le potentiel de contact passe à $+0,6$ V dans les mêmes conditions, soit une différence de l'ordre du volt.

Avec une structure à gate d'aluminium, on avait calculé :

$$V_{TH} = -1,05 - 1,4 - 0,35 - 0,5 = -3,3 \text{ V}$$

Si l'on passe au gate au silicium, on trouvera dans les mêmes conditions :

$$V_{TH} = -1,05 - 1,4 + 0,6 - 0,5 = -2,35 \text{ V}$$

En dopant diversement ce silicium, on agit d'ailleurs sur la tension de seuil qui peut être ajustée à volonté.

Les premiers transistors MOS à gate au silicium (les *silicon gate MOS* des Américains) sont dus aux travaux, de J. C. SARACE et de son équipe, aux *Bell Telephone Laboratories*.

Le processus de fabrication d'un MOS à gate en silicium est développé dans la figure 3, selon la méthode élaborée dès 1967 par *Fairchild*. On part d'un substrat N sur lequel on fait croître une couche épaisse d'oxyde : puis on ouvre des fenêtres, une par transistor à fabriquer, et l'on remplace l'oxyde épais par de l'oxyde mince : c'est là qu'intervient donc le premier masquage (en *a*).

Après quoi, toute la surface du wafer (du disque de silicium sur lequel sont réalisés les CI) est recouverte de silicium polycristallin (*b*). Un second masquage délimite la zone de gate, et par la même occasion, un éventuel premier plan d'interconnexions. L'oxyde mince et le silicium polycristallin sont éliminés, sauf à l'emplacement du gate, comme le montre la figure 3 *c* : cet emplacement circonscrit très précisément l'action de la diffusion du bore qui, sans masquage supplémentaire, est immédiatement assurée (*d*). Le silicium polycristallin du gate est donc, lui aussi, dopé à ce moment, et sa résistivité tombe à $1 \Omega/\text{cm}$ environ.

Une première différence importante apparaît donc à ce niveau avec les structures classiques : c'est ce que l'on a voulu préciser figure 4. On voit que le chevauchement du gate en aluminium sur le drain et la source a disparu avec la technique du gate en silicium, d'où il s'ensuit une réduction notable des capacités.

Mais revenons au processus technologique. Le bore, qui diffuse relativement vite dans le silicium, laisse pratiquement intact l'oxyde de silicium ; cette diffusion (la seule dans tout le processus de fabrication) vise non seulement à créer le drain et la source, mais également à doper le gate pour l'amener à une faible résistivité, d'environ $1 \Omega/\text{cm}$.

Suivent une nouvelle oxydation et un masquage, le troisième, qui prépare le dégagement des prises de contact de drain et de source (*e*). Une couche d'aluminium est alors déposée sur toute la surface du wafer, un quatrième et dernier masquage intervenant pour n'en laisser subsister que ce qui est nécessaire au second plan d'interconnexions (*f*).

La technologie à gate au silicium s'est rapidement développée en raison de son intérêt. On notera d'ailleurs qu'elle est compatible avec d'autres processus technologiques, comme on l'évoquera ultérieurement. Notons à son actif :

1. *La notion de rendement en fabrication*, qui est essentielle en microélectronique, intervient en tout premier lieu : or, les rendements sont, ici, supérieurs à ceux obtenus avec la technologie fondamentale parce que l'oxyde du gate, point sensible de l'élément, est aussitôt recouvert par du silicium qui le protège de toute contamination ultérieure.

2. *On aura remarqué qu'avec ce processus, le « chevauchement » du gate sur les zones de source et de drain est, sinon totalement supprimé, du moins réduit au minimum* que permet la diffusion.

En effet, celle-ci est assurée après réalisation du gate : *on dit que ce processus est à auto-alignement des électrodes*. Le recouvrement peut alors être réduit jusqu'au micromètre, alors qu'il est normalement de 5 à 10 μm avec la technologie de base.

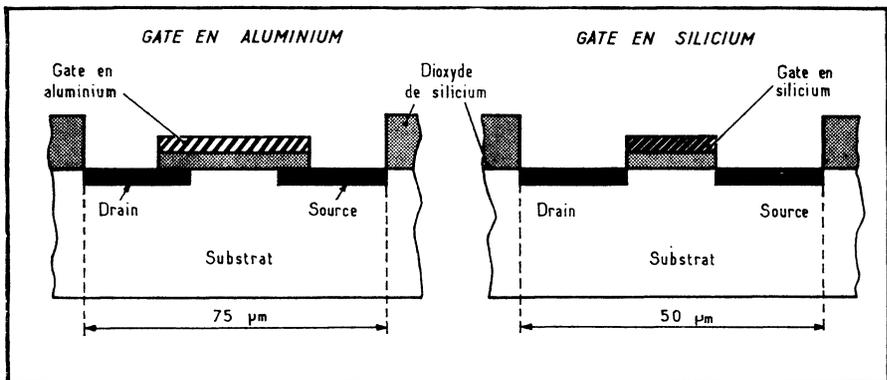


Fig. 4. — Structures comparées d'un MOS classique et d'un MOS à gate en silicium. Les électrodes de la structure CIS (« *conducteur-isolant-semiconducteur* ») sont auto-alignées ce qui évite tout chevauchement du gate sur le drain et la source, donc réduit considérablement les capacités de gate.

3. *Le gate peut, de plus, être plus court*. Ces deux raisons contribuent à une réduction des capacités du transistor, donc à un accroissement du gain et, ce qui est encore plus important, les vitesses de travail augmentent sensiblement.

4. *La tension de seuil est abaissée à environ 2 V*. On peut à nouveau réduire les tensions d'alimentation et les rendre, ainsi que les niveaux logiques, compatibles avec les logiques bipolaires DTL/TTL.

5. *On a vu que le procédé de fabrication fournit deux couches d'interconnexions* : c'est là une qualité des plus appréciables en LSI. Si l'on inclut la couche d'interconnexions possible due au substrat, on aboutit à un total de 3 couches.

6. Enfin, puisque l'oxyde de gate est immédiatement protégé, on peut faire subir au wafer des traitements ultérieurs à haute température sans risque, et, éventuellement, mêler des transistors bipolaires et MOS.

On obtient les caractéristiques générales suivantes :

$$V_{TH} = -2 \text{ à } -3 \text{ V}$$

$$V_{TF} = -30 \text{ à } -50 \text{ V}$$

$$V_{DD} = +5 \text{ V}$$

$$V_{GG} = -12 \text{ V}$$

D'autre part, la surface de silicium occupée est moindre, donc la densité d'intégration s'accroît; par contre, le processus de fabrication est un peu plus complexe qu'avec la technologie standard.

Le tableau B résume les caractéristiques des quatre techniques qui viennent d'être présentées.

Tableau B. — Les 4 techniques fondamentales

CARACTÉRISTIQUES	MOS P <111>	Sur Si <100>	Au nitrure de Si	Gate au silicium
Tension de seuil V_{TH} (V)	-3 à -5	-1,7 à -2,2	-2 à -2,5	-2 à -3
Tension d'inversion V_{TF} (V)	-30 à -40	-15 à -20	-30 à -50	-30 à -50
Densité relative d'intégration	0,7	0,6	0,8	1
Tensions d'alimentation types (V)				
• V_{DD}	-13	+5	+5	+5
• V_{GG}	-27	-12	-12	-12

6. LE MOS A CANAL N

Bien que ne constituant pas à proprement parler une variante technologique, il convient de revenir un instant sur le NMOS, ou MOS à canal N.

Ici, la fabrication part d'un substrat P, dopé au bore, par exemple. La tension de seuil est très faible, et positive, de l'ordre du volt. Si l'on considère un NMOS à gate au silicium, on rencontrera les étapes suivantes de fabrication (fig. 5) :

- on fait croître de l'oxyde thermique sur environ 1 000 nm;
- le 1^{er} masquage prépare la fabrication du gate;
- le diélectrique de gate et le silicium polycristallin de gate sont déposés;
- le 2^e masquage prépare la diffusion N;
- on exécute la diffusion N;
- un dépôt de verre précède le 3^e masquage qui ouvre le passage aux connexions;
- de l'aluminium est évaporé et gravé, grâce au 4^e masquage;
- le 5^e masquage suit un dépôt de verre servant à la passivation et permet la sortie des connexions.

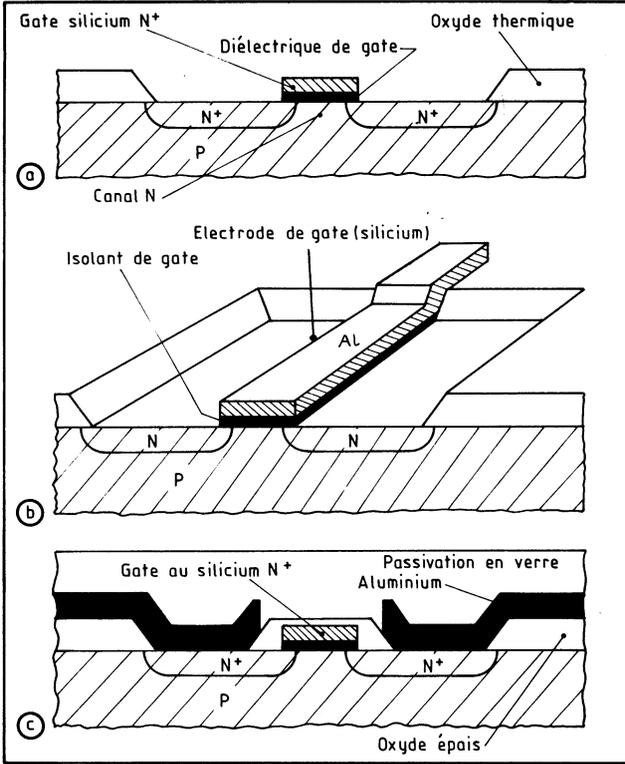


Fig. 5. — Fabrication du MOS à canal N.

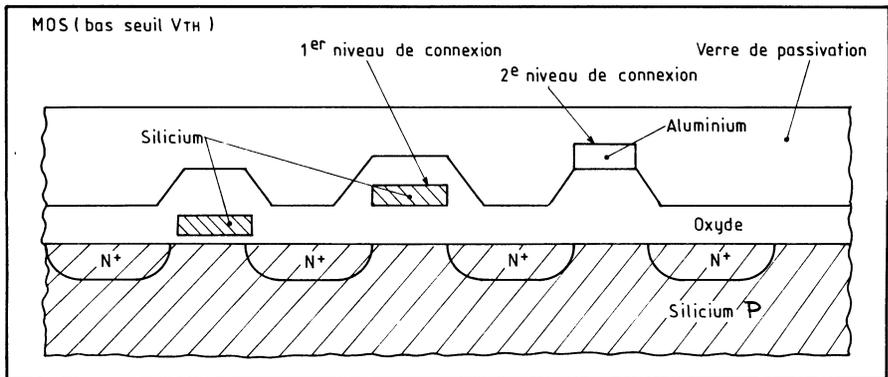
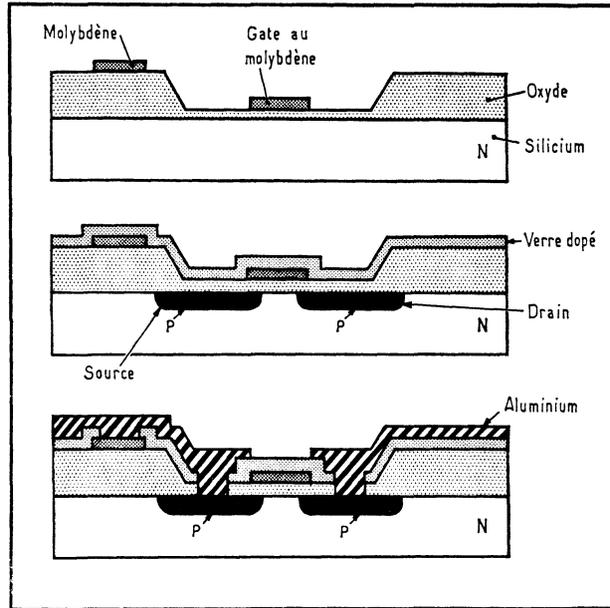


Fig. 6. — Les deux niveaux d'interconnexion du NMOS (auxquels s'ajoutent les connexions possibles via le silicium).

Fig. 7. — Trois étapes importantes dans la fabrication d'un MOS à gate au molybdène dit RMOS. (Selon *General Electric*.)



En raison d'une tension de seuil très basse, ce qui est à la fois un avantage et un gêne, on peut utiliser des tensions d'alimentation très faibles et parfaitement compatibles TTL, une tension unique de +5 V par exemple.

Les dimensions du MOS dépendant également des tensions appliquées, on peut les réduire notablement et, par exemple, rapprocher les diffusions à 10 μm ou réaliser un canal de 4 à 5 μm de longueur. Il en résulte une vitesse encore accrue, car aux faibles dimensions s'ajoute la mobilité supérieure des porteurs ($\mu_N = 2$ à 3 fois $\mu_P = 400 \text{ cm}^2/(\text{Vs})$, environ). La résistance du canal $R_{DS(on)}$ est réduite de moitié, et les fréquences de fonctionnement doublent ou triplent. D'autre part, la densité d'intégration est environ le double de celle des PMOS.

En revanche, la tension d'inversion est faible, et le procédé peut exiger l'utilisation d'anneaux de garde de type P pour limiter l'importance des champs électriques. La figure 6 montre, en coupe, où se situent les niveaux de connexions.

7. LE MOS A GATE AU MOLYBDÈNE (RMOS)

Variante de la structure à gate au silicium, le RMOS (R venant de *réfractaire*) était proposé vers 1968 par les D^{rs} Dale M. BROWN, William ENGELER, Marvin GARFINKEL et Peter V. GRAY, du centre de recherche et développement de *General Electric*, puis réalisé en 1970.

En 1971, la firme proposait une série de circuits RMOS fonctionnant typiquement à 5 MHz, dont un double registre 100 bits à décalage dynamique (GER 1507

et GER 2507), une RAM hybride 4 096 bits à cycle de 125 ns (GER 1036) dont l'élément de base est de 256 bits, tous circuits compatibles TTL. Il ne semble pas qu'ils aient connu de développement ultérieur. Par contre, d'autres centres de recherche ont étudié, entre temps, le RMOS, et parmi eux le LEP français qui a réalisé des dispositifs expérimentaux.

Le MOS utilise un métal à point de fusion élevé (et appelé peut-être abusivement « métal réfractaire »), tel que le tungstène, ou surtout le molybdène (2 600 °C); ce dernier est bien meilleur conducteur que le silicium (le gain en résistance serait de l'ordre de 200).

Dans la technologie RMOS, le gate au molybdène procure un auto-alignement des électrodes (comme avec le gate au silicium) et deux niveaux d'interconnexions. Le processus de fabrication est le suivant.

On part d'un substrat de silicium N, que l'on oxyde avec un film de $1,3 \mu\text{m}$ d'épaisseur. Des fenêtres sont ouvertes puis refermées avec une couche mince d'oxyde ($0,1 \mu\text{m}$). Le molybdène est déposé sur toute la surface, puis ôté sauf à l'emplacement du gate et du premier niveau d'interconnexions (fig. 7).

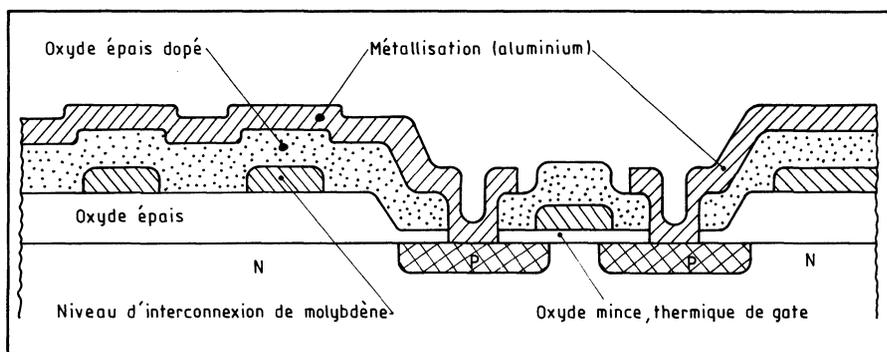


Fig. 8. — Le double niveau d'interconnexion du RMOS.

Ensuite, un verre dopé au bore est déposé sur l'ensemble et une diffusion est assurée (vers $1\ 100\text{ °C}$). Le bore diffuse à travers la couche mince d'oxyde, la couche épaisse comme le molybdène constituant un barrage infranchissable. Ainsi sont formés les drain et source : le molybdène a donc servi à auto-aligner les électrodes et leur recouvrement (leur chevauchement) reste réduit au minimum. Enfin, les contacts sont ouverts et les interconnexions (second niveau) assurées avec de l'aluminium. Quatre masquages auront été nécessaires.

Avec le RMOS, la tension de seuil est également ramenée à 1,5 à 2,5 V pour une densité d'intégration accrue comme on le conçoit avec ce processus à auto-alignement.

D'autres chiffres étaient fournis par GE :

- la résistance superficielle de la région diffusée P est de $10 \Omega/\text{carré}$ (contre $30 \Omega/\text{carré}$ pour le MOS classique, toujours sur substrat orienté $\langle 1.1.1. \rangle$);
- la résistance superficielle du gate est de $0,15 \Omega/\text{carré}$ contre $40 \Omega/\text{carré}$ pour le gate au silicium et $0,05 \Omega/\text{carré}$ pour le gate en aluminium).

Les RMOS sont, bien sûr, compatibles DTL/TTL. L'obtention d'un double niveau de connexions est illustré, lui, par la figure 8 où l'on voit que des connexions en molybdène, isolées de l'aluminium, ont été réalisées.

Un autre avantage du molybdène, c'est son inertie chimique à l'égard de la silice. D'autre part, il n'est pas attaqué par les réactifs habituels utilisés pour graver la silice. Il semble enfin que le molybdène améliore la fiabilité des dispositifs.

Le *LEP*, lui, a utilisé cette technologie conjointement avec l'implantation ionique, selon un processus développé par M. BERTH et présenté à l'occasion d'une demi-journée d'étude SEE (24 janvier 1973). Le *LEP* utilise des plaquettes de silicium N, de 5 à 7 Ω/cm , en $\langle 111 \rangle$.

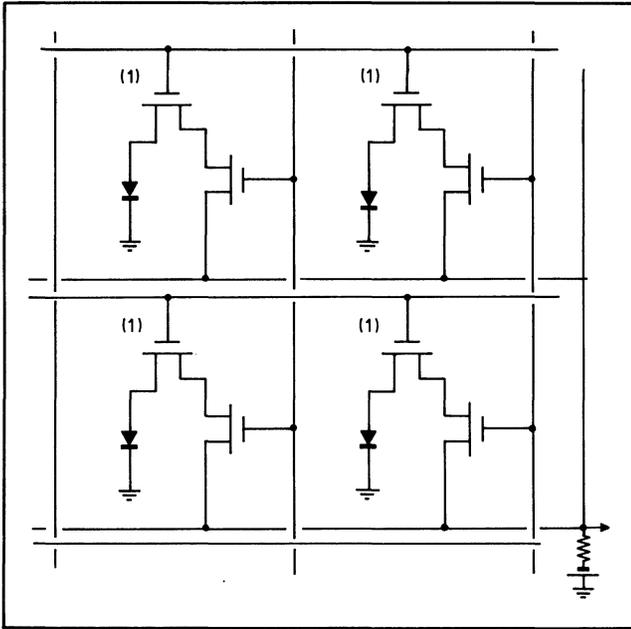


Fig. 9. — Principe du dispositif expérimental de prise de vues à RMOS du LEP.

Ce RMOS a servi, au *LEP*, à réaliser un dispositif expérimental de prise de vues à état solide dont chaque élément correspondant à un point d'image est constitué par deux transistors MOS connectés en série (fig. 9). En appliquant des séquences d'impulsions de tension aux lignes et aux colonnes du « cross-bar » d'adressage, on met successivement en relation avec la source de tension les jonctions source-substrat (qui jouent le rôle d'éléments photosensibles) des MOS «1»; on restitue ainsi à chaque photodiode la charge qu'elle a perdue par effet photoélectrique depuis la consultation précédente; le train d'impulsions qui apparaît aux bornes de la résistance de charge constitue le signal vidéo.

8. MOS A DOUBLE DIFFUSION (D/MOS)

Le processus à double diffusion « D/MOS » a été mis au point par *Signetics* et annoncé en décembre 1970. Il recourt à des MOS à canal N et s'appliquerait aussi bien aux CI logiques qu'aux éléments discrets jusqu'aux micro-ondes : une fréquence de 10 GHz (gigahertz) aurait été obtenue en laboratoire.

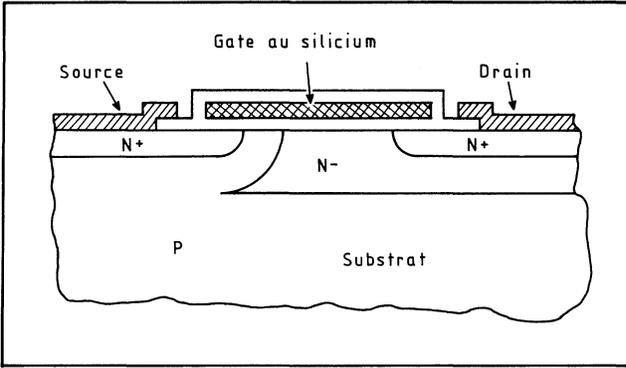


Fig. 10. — Structure d'un DMOS *Signetics*, MOS à canal N à double diffusion, qui atteindrait 10 GHz, soit un temps de montée de 0,2 ns (210 picosecondes exactement!).

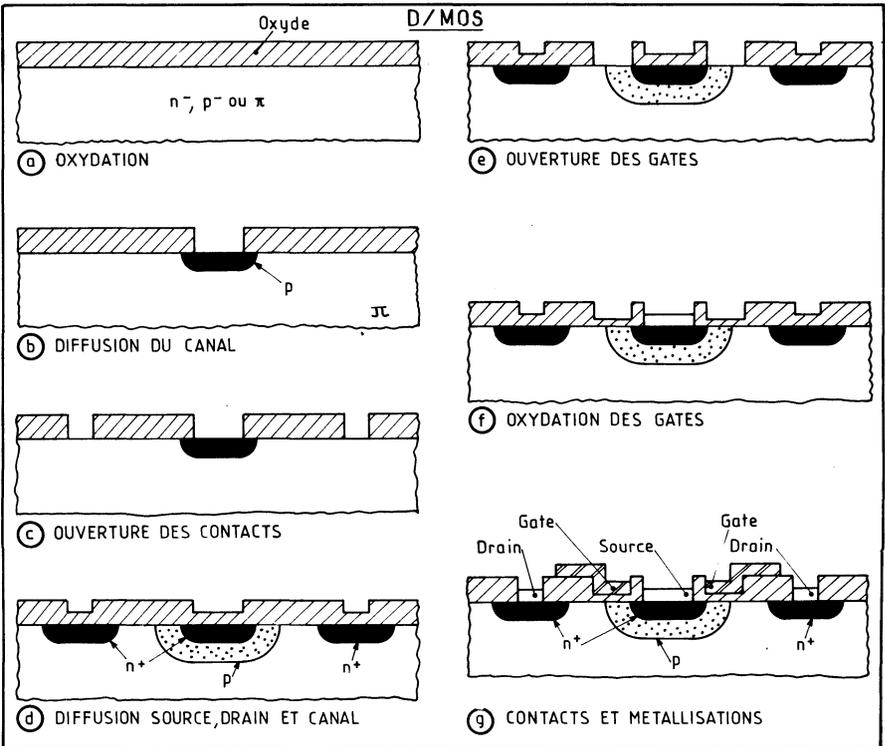


Fig. 11. — Fabrication du DMOS.

Avec les D/MOS, on crée aussi bien des éléments à déplétion qu'à enrichissement et leur vitesse est d'environ 5 fois supérieure à celle des MOS classiques N. Cela est dû au fait que la longueur du canal a été réduite (donc, le temps de transit est réduit) grâce à une astuce de double diffusion d'où l'appellation de D/MOS (fig. 10).

La longueur du canal est, en effet, la différence entre les diffusions P et N +, et est ramenée à environ 1,5 micron (contre les 5 μm habituels). L'élément est fabriqué sur un substrat N — épitaxial à partir de : deux diffusions P et N réalisées à travers les mêmes fenêtres de l'oxyde du côté source, une seule diffusion du côté drain. Le dopage P détermine la tension de seuil et le circuit obtenu peut travailler sous 5 V ou même moins (fig. 11).

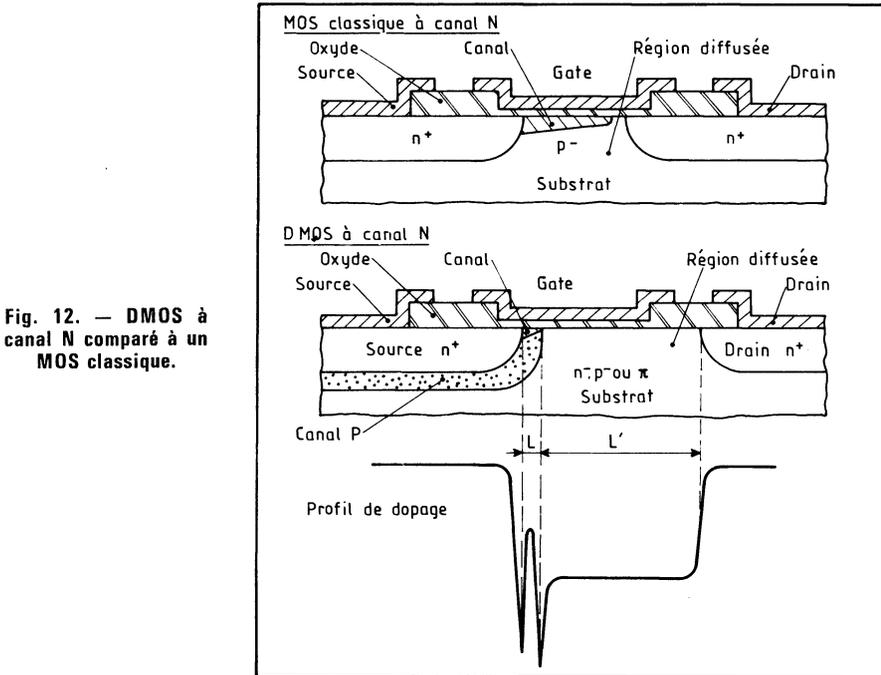


Fig. 12. — DMOS à canal N comparé à un MOS classique.

En fait, le substrat peut être N—, ou P—, ou même intrinsèque, comme le montre la figure 12 qui, avec son profil de dopage, compare le D/MOS à un MOS classique canal N. A la limite, on pourrait assimiler le canal P à la base P d'un transistor bipolaire NPN. Enfin, la technologie à gate au silicium s'applique également au D/MOS (fig. 13).

Outre sa fréquence de travail qui peut dépasser le gigahertz, le D/MOS offre une méthode simple pour réduire la résistance à l'état passant R_{on} du MOS. Celle-ci varie normalement entre 100 et 500 Ω et dépend, en particulier, du rapport W/L du canal. Ainsi, avec un MOS classique à canal $L = 8 \mu\text{m}$, il faudrait un rapport $W/L = 800$ environ pour aboutir à une $R_{DS(on)}$ de 15 Ω , soit $W = 6,4 \text{ mm}$, ce qui est bien évidemment prohibitif; la surface d'un tel MOS atteindrait 0,2 mm^2 .

Par contre, en D/MOS et avec le même rapport de 800, il suffit d'une largeur $W = 800 \mu\text{m}$ pour $L = 1 \mu\text{m}$, soit une surface totale, pour le MOS, de $0,025 \text{ mm}^2$.

D'autre part, et en raison de ses faibles dimensions, les capacités du D/MOS sont réduites et les temps de commutation peuvent être très courts, 600 ps par exemple. Ces faibles capacités réduisent la diaphonie; les pertes d'insertion étant également réduites, on comprend que les premières applications des D/MOS, en 1975 se soient orientées vers la commutation téléphonique.

9. LES CMOS, OU MOS COMPLÉMENTAIRES

Sans pour autant développer ici les caractéristiques des CMOS, auxquels un chapitre est consacré, il faut cependant noter leur technologie. En particulier, parce qu'on peut revenir, avec eux, à un substrat orienté $\langle 100 \rangle$.

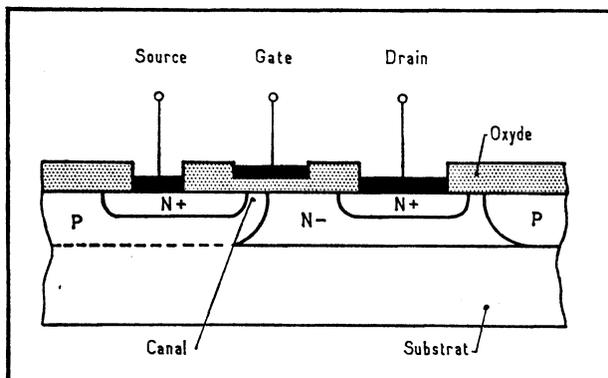


Fig. 13. — Le gate au silicium confère l'auto-alignement au DMOS, mais cet avantage n'est utile que du côté de la source.

La structure du CMOS, ou plutôt d'une paire de MOS complémentaires montée en inverseur fondamental, est donnée figure 14. La première opération de fabrication consiste à diffuser une zone P qui servira de caisson au NMOS; elle peut encore être réalisée par implantation ionique. Une diffusion P+ crée ensuite les source et drain du MOS à canal P et les diodes de protection (fig. 15).

De la même façon, une diffusion N+ donne naissance aux drain et source de l'élément à canal N et sert à compléter le système de protection.

Pour éviter les contaminations par ions + si redoutables, comme on l'a vu, pour les MOS à canal N, une technologie d'oxyde pur a dû être élaborée.

D'autre part, des anneaux de garde entourent totalement les composants pour prévenir toute fuite éventuelle. Pour les MOS à canal P, ce sont des zones diffusées N+ qui servent également à amener la tension d'alimentation V_{DD} à travers le substrat N aux MOS à canal N reliés à V_{DD} . Pour les MOS à canal N, ce sont des zones diffusées P+ qui servent également de conducteur au courant de masse (V_{SS}) appliqué aux MOS à canal N reliés à la masse, à travers les zones P. La métallisation est ensuite assurée de façon habituelle.

Fig. 14. — Principe des MOS complémentaires (CMOS).

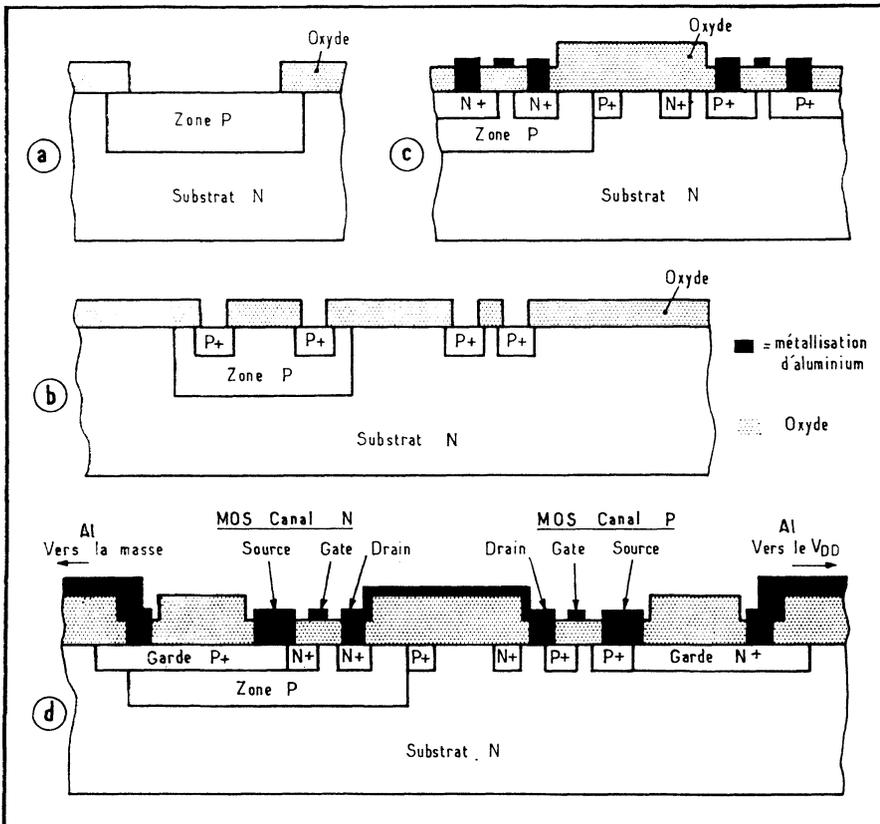
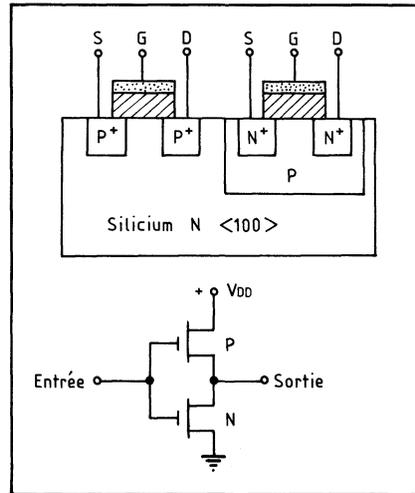


Fig. 15. — Fabrication d'un inverseur CMOS. On crée une zone P où sera réalisé le MOS à canal N (a); cette zone reçoit une diffusion P+ (b). La paire complémentaire apparaît en cours de fabrication en (c), et est achevée en (d).

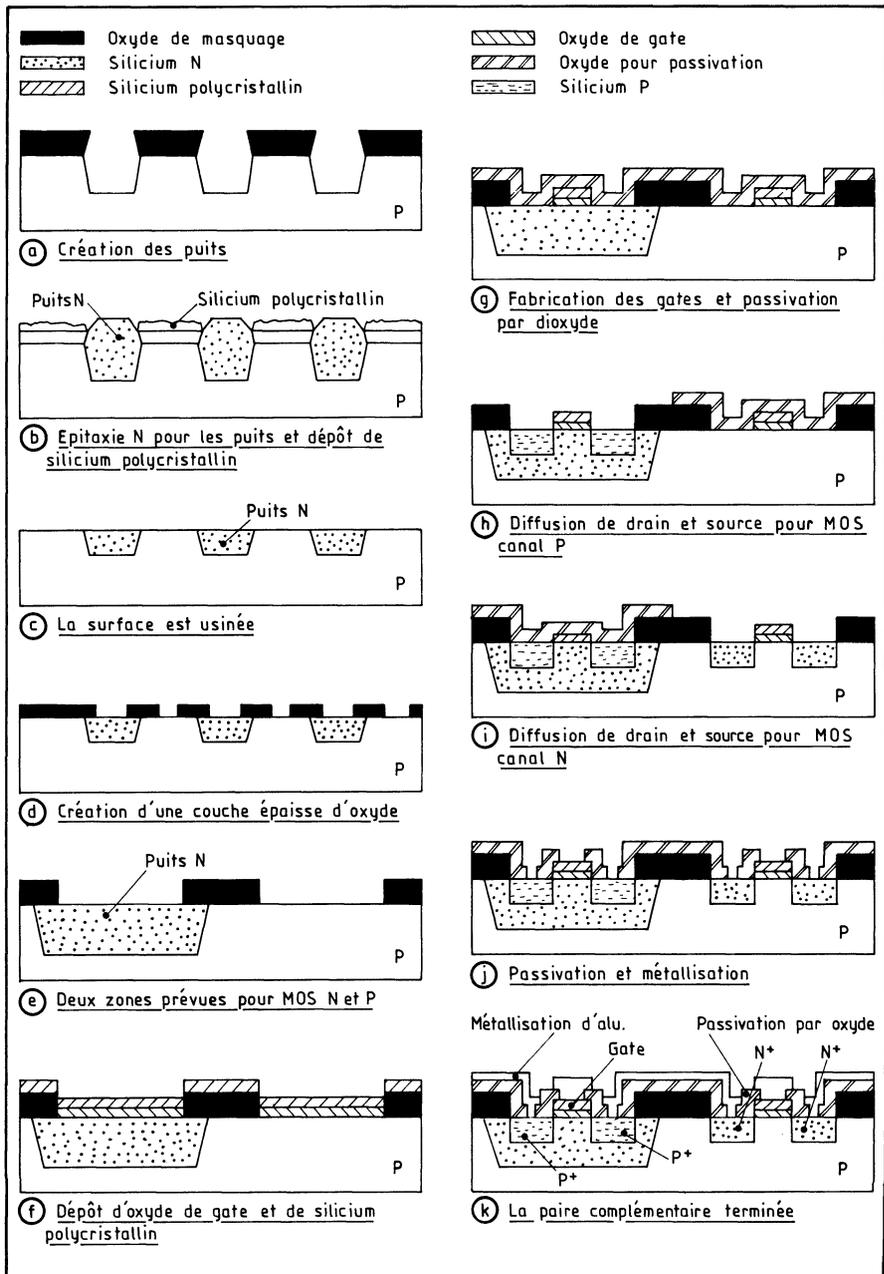


Fig. 16. — Processus de fabrication des CMOS à gate au silicium selon Motorola.

CMOS à gate au silicium

Les combinaisons de technologies peuvent prendre toutes les formes théoriquement possibles. Cependant, certaines formules semblent plus attrayantes que d'autres et par exemple, celle qui combine les MOS complémentaires avec le gate au silicium.

Une telle formule, proposée par *Motorola* en 1971 pour des applications à l'horlogerie électronique, est en effet intéressante si l'on veut réduire la consommation, ou abaisser les tensions d'alimentation. Sa fabrication est brossée dans le film de la figure 16.

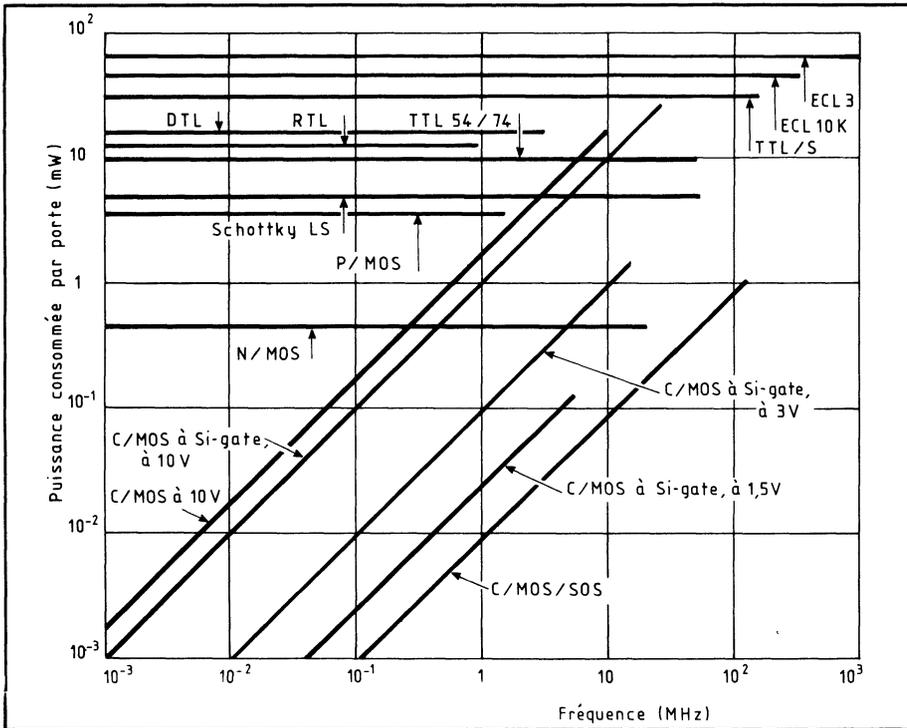


Fig. 17. — Cette famille de courbes montre comment varie la consommation avec la fréquence pour quelques familles de CI bipolaires et MOS. On note l'avantage évident en faveur des MOS complémentaires à gate au silicium.

Dans un substrat P à orientation cristalline $\langle 1.0.0 \rangle$ sont créés des « puits », par usinage chimique, remplis par épitaxie de silicium N. Du silicium polycristallin recouvrant l'oxyde de masquage permet ensuite le polissage de la surface générale du wafer. Une oxydation épaisse (sur 0,5 à 1 μm) est alors réalisée, et à ce moment commence la fabrication des composants proprement dite.

A cet effet, on fait croître l'oxyde, qui sera retenu pour les gates des éléments N et P sur une épaisseur de 1 000 Å environ, puis on recouvre le tout de silicium polycristallin. La structure des MOS, P et N, se dessine à ce moment et le processus qui suit, avec les diffusions de gate et de drain (N ou P selon le canal des MOS à produire) est classique.

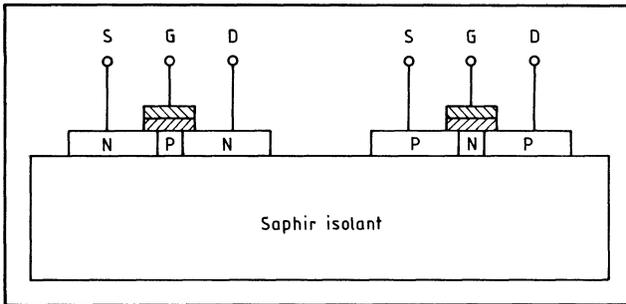


Fig. 18. — Structure CMOS sur saphir (SOS)

La figure 17 précise d'ores et déjà les consommations, en fonction de la fréquence, pour les portes fondamentales des principales familles MOS et bipolaires. On reviendra plus longuement sur les CMOS ultérieurement, mais il convient d'évoquer ici leur prolongement en SOS.

10. CIRCUITS SUR SUBSTRAT ISOLANT : LE SOS

Tous les CI produits, en monolithique, sont réalisés sur un substrat semi-conducteur. Or, les caractéristiques obtenues seraient évidemment meilleures si le substrat était isolant : absence de courants de fuite, d'éléments parasites...

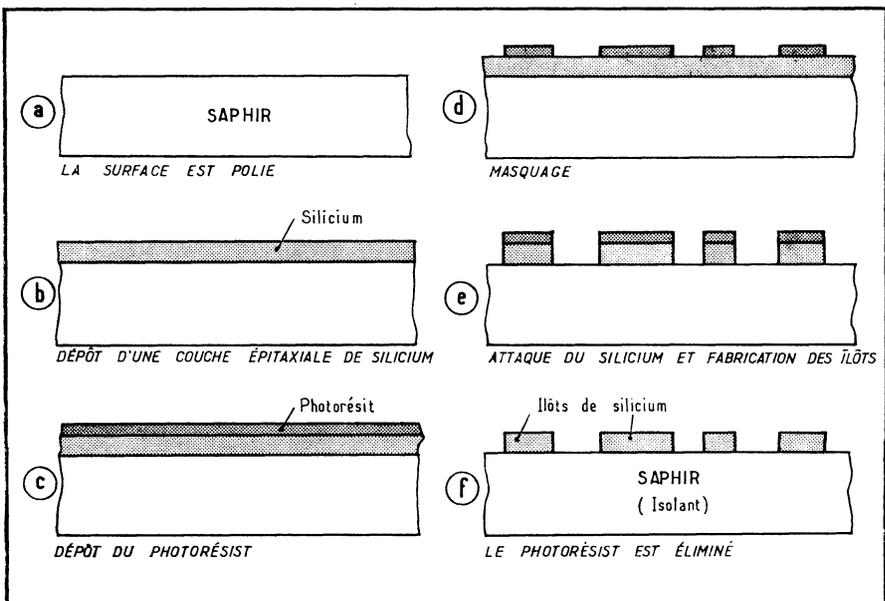


Fig. 19. — Les étapes de fabrication d'un CI sur substrat isolant, saphir ou spinelle. Après l'étape (f) commencera la production des composants proprement dits, dans les îlots réservés à cet effet, et selon le processus habituel.

On a donc songé à remplacer l'habituel silicium (semiconducteur) par un matériau isolant dont le réseau cristallin ressemble le plus à celui du silicium : ce sera soit *le saphir*, variété bleue de corindon, soit *le spinelle* qui est un aluminat naturel du magnésium $MgAl_2O_4$. C'est pourquoi cette technologie porte le nom de SOS (*Silicon On Sapphire* ou *Silicon On Spinel*).

Cette technologie convient à merveille aux CMOS, dont elle fait les produits les plus exceptionnels qui soient : économie maximale de puissance, très haute rapidité (t_{pd} vers les 3 ns dès 1975). Une structure CMOS/SOS est ainsi stylisée par la figure 18.

Le processus de fabrication illustré figure 19 montre que, à partir du dépôt de la couche épitaxiale, on crée des îlots de silicium semiconducteur parfaitement isolés les uns des autres. C'est dans ces îlots que seront ensuite produits les composants, cette fois de façon traditionnelle.

RCA a œuvré en tant que pionnier dans cette technologie; en France, EFCIS produit des circuits en tous points remarquables en SOS depuis plusieurs années déjà. Un autre fabricant s'est également lancé dans cette voie, mais pour ses propres besoins : *Hewlett-Packard*, pour ses matériels d'instrumentation et informatique. C'est qu'en effet, les CMOS/SOS consomment excessivement peu tout en procurant des vitesses élevées : 3 ns pour un inverseur ne consommant que 5 nW, par exemple, soit un facteur de qualité de 15 millièmes de picojoules au repos! Leur immunité au bruit est, d'autre part, excellente : c'est celle des CMOS.

Ce qui a freiné leur développement, c'est le prix du Saphir. Or, RCA (ainsi que d'autres laboratoires de recherches) a proposé de remplacer la fabrication du saphir en lingots débité ensuite en tranches (avec près de 45 % de pertes dues au sciage!) par une fabrication en ruban, qui en réduirait le prix du tiers (1). Ainsi, le SOS deviendrait compétitif avec le silicium.

Il a même été proposé de « noyer » les composants CMOS dans le substrat de saphir afin d'aboutir à une structure plane, plus facile à traiter et probablement plus fiable. Selon cette méthode appelée SIS (pour « *silicon in sapphire* »), le saphir est usiné par bombardement ionique jusqu'à l'obtention de trous de $0,6 \mu m$ de profondeur, délimités bien sûr par un masque et obtenus en une heure d'usinage. Cette méthode a été préconisée par C. E. WEITZEL, des laboratoires RCA de Princeton. Une fois les ouvertures pratiquées dans le saphir, on recouvre toute sa surface de silicium épitaxial qui sera usiné à son tour, afin qu'il ne subsiste que dans les puits créés et arrive ainsi à fleur de surface.

11. TECHNIQUES A OXYDE DANS LE SILICIUM

La surface tourmentée des CI met en relief les angles vifs qui créent des points fragiles pour la métallisation (fig. 20). Pour éviter ces points éventuels de rupture, plusieurs techniques ont été élaborées, dont la LOCOS de *Philips*, la *Planox* de *SGS*, le *SATO* de *Texas*, l'*Isoplanar* de *Fairchild*, etc. : elles visent à ensevelir partiellement dans le silicium la couche d'oxyde épais, et ainsi, à rendre le relief en surface plus régulier, comme le montre la technique LOCOS.

1. Voir la revue « *Électronique et Applications Industrielles* », n° 249, du 15-3-78, p. 21 : *Le procédé EFG*, par B. V. VONDERSCHMITT, vice-président de RCA Corps.

De ce fait, les capacités parasites entre MOS sont réduites, ainsi que les courants de fuites; la tension d'inversion au-dessus de cet oxyde est fortement accrue; de plus, de telles techniques sont souvent génératrices de MOS auto-alignés, ce qui permet de réduire les dimensions des transistors et, partant, d'accroître la densité d'intégration et la vitesse de fonctionnement.

Élaborée par *Philips*, la technique LOCOS (*local oxidation of silicon*), a d'ailleurs été développée en LOCOS pour les MOS complémentaires, et même en versions pour familles bipolaires ultrarapides (ECL subnanosecondes).

12. L'ISOLEMENT DIÉLECTRIQUE

Le thyristor parasite des CMOS

Ainsi qu'on l'a vu, les CMOS doivent subir une diffusion supplémentaire qui crée un caisson P pour les MOS à canal N (dans le cas où l'on ne recourt pas au SOS), ainsi que des diffusions pour les anneaux de garde.

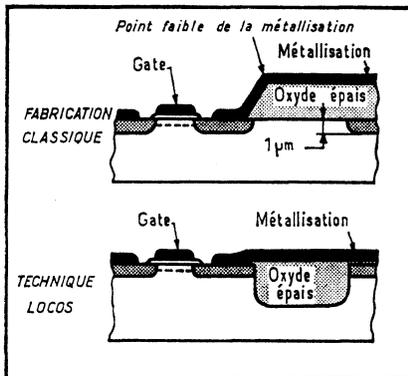


Fig. 20. — Avec la technique LOCOS, Philips « enterre » la couche épaisse d'oxyde et évite ainsi les angles vifs et les risques de rupture de la métallisation.

Or, l'isolement des MOS est tout relatif puisqu'il est alors assuré, comme pour les bipolaires d'ailleurs, par des jonctions bloquées, polarisées en inverse; ainsi en va-t-il pour les sources et drains, par rapport au substrat. Si, alors, une tension supérieure à celle du substrat est appliquée à ces électrodes, un courant passe dans ce que l'on peut assimiler à un transistor bipolaire équivalent (fig. 22). Ce cas est plus probable lorsque le circuit est utilisé en commutateur de tensions analogiques et que les lignes sont assez longues et introduisent des transitoires ou du bruit.

La situation devient encore plus désastreuse lorsque les zones de substrat P— et N— ne sont plus référencées à leur potentiel, par accident. Alors, on se trouve en présence de deux transistors bipolaires rebouclés, fonctionnant en thyristor, c'est-à-dire en avalanche (fig. 23). Il en résulte, si le signal est présent aux sources et drains, la destruction pure et simple de l'élément.

Pour éviter de tels accidents, on connecte généralement des résistances en série avec tous les accès au circuit; elles limiteront le courant à un niveau acceptable en cas de déclenchement intempestif en avalanche.

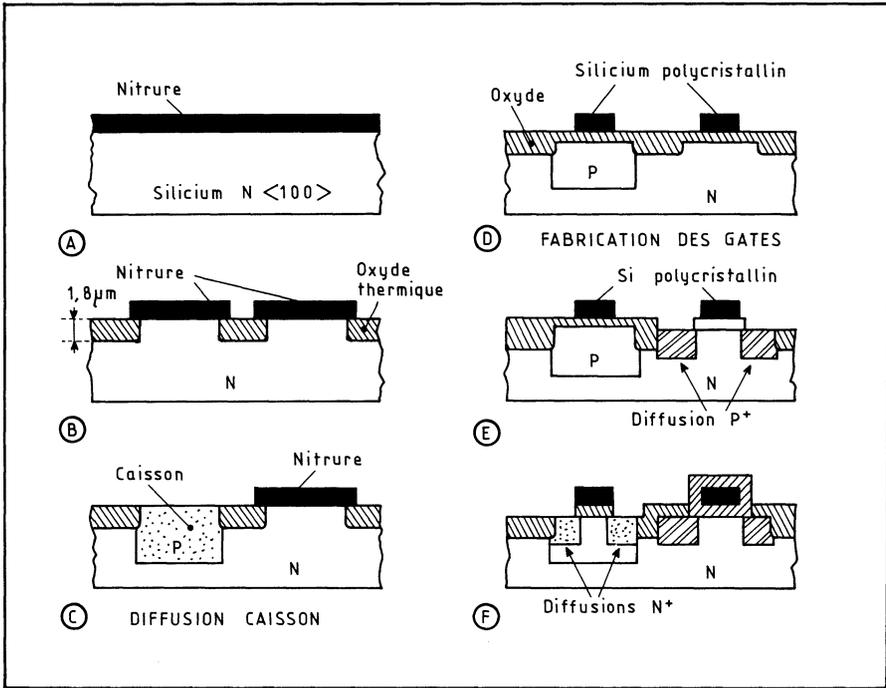


Fig. 21. — Principe du LOCOS : c'est la technique LOCOS appliquée aux MOS complémentaires. Elle s'applique également au bipolaire et, pour l'ECL, donne la technologie « ILO », et même « SUB-ILO » pour les ECL subnanosecondes.

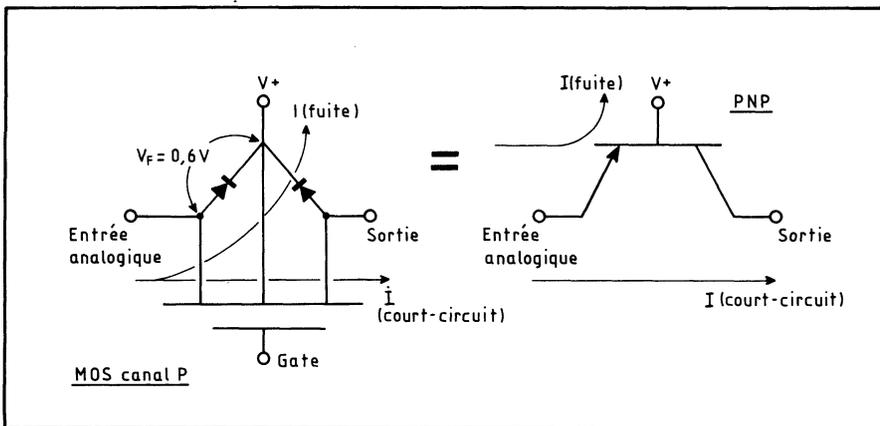


Fig. 22. — Le transistor parasite du MOS.

Une autre méthode, comparable dans son effet à la technologie CMOS/SOS, mais recourant, elle, à un substrat de silicium, est concevable : c'est la technique d'isolement diélectrique dont *Harris Semiconductor* s'est fait le champion. Grâce à elle, tous les composants sont isolés par du dioxyde de silicium, et par conséquent tous les éléments parasites sont supprimés (fig. 24); en effet, elle aboutit à constituer de véritables cuvettes de SiO_2 .

La technologie EPIC

Cette technologie, encore connue sous le nom d'EPIC (dénomination *Motorola* : « *epitaxial passivated integrated circuit* »), comporte les étapes suivantes (fig. 25).

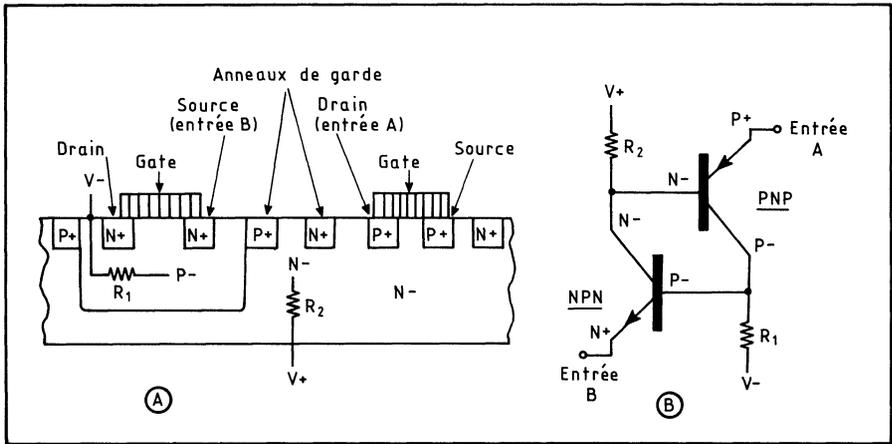


Fig. 23. — Le thyristor parasite des CMOS.

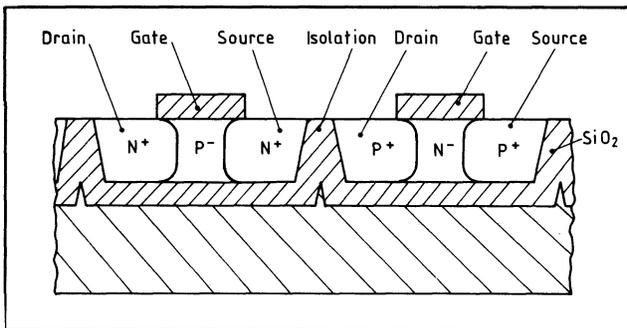


Fig. 24. — L'isolement diélectrique EPIC.

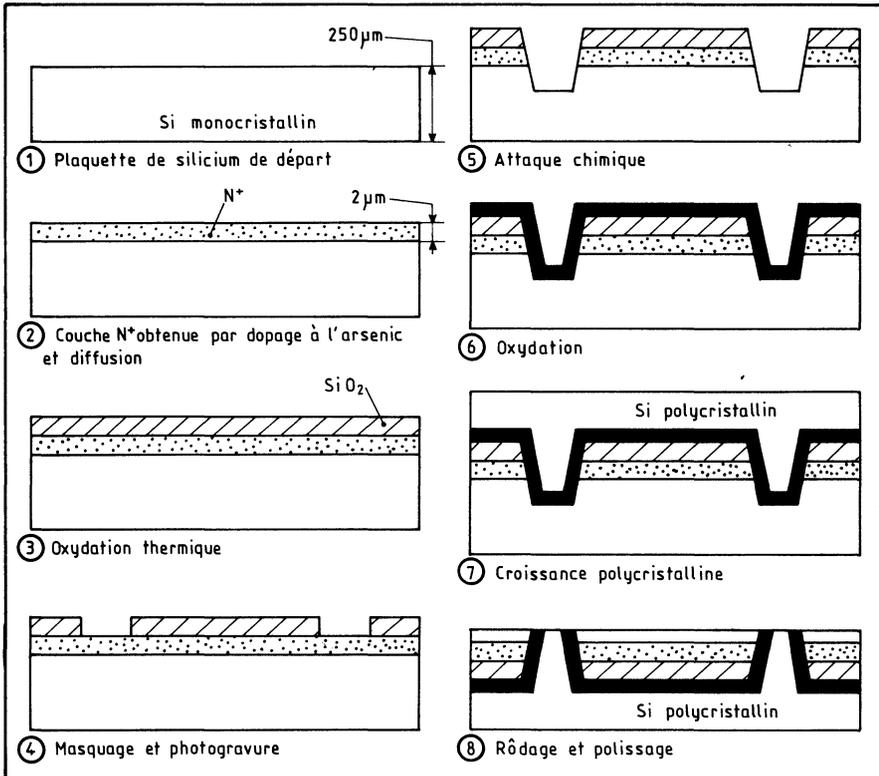


Fig. 25. — Fabrication d'une cuvette de silice.

On part d'un substrat de type n , sur lequel on diffuse préalablement une couche n^+ ; puis, sa surface est oxydée, et des fenêtres sont ouvertes là où se trouveront les parois des caissons. Le silicium est ensuite attaqué, de façon à dégager des « sillons », dont la surface est également oxydée. Ces sillons sont alors remplis de silicium polycristallin cette fois, dont une couche recouvre toute la surface du substrat. La face inférieure de ce dernier est alors usinée, jusqu'à effleurement du dos des sillons, après quoi l'ensemble est retourné et se prête, dès lors, au traitement classique de réalisation des circuits intégrés.

Cet isolement diélectrique (d'où son sigle « DI ») a d'ailleurs été d'abord appliqué aux circuits bipolaires. Avec les CMOS, il offre quasiment les mêmes avantages que le SOS en vitesse et consommation; c'est cependant une technologie difficile à maîtriser.

13. MOS A COUCHE MINCE SUR SUBSTRAT ISOLANT

Il est encore possible de produire des MOS sur un substrat isolant selon la technologie préconisée, par exemple, par *Philips*.

Elle consiste à utiliser un substrat de silicium qui sera ensuite éliminé par une attaque électrochimique sélective en fonction du degré de dopage du substrat : l'épaisseur du silicium conservé est d'environ $0,5 \mu\text{m}$. Le MOS ainsi réalisé — ou le circuit intégré — est disponible sur une couche isolante, soit d'oxyde de silicium, soit de céramique.

La figure 26 montre la méthode grâce à laquelle la couche isolante est appliquée avant l'attaque. Sur la couche épitaxiale, on fait croître une pellicule d'oxyde de $1 \mu\text{m}$ environ et on fait évaporer sur cette pellicule très mince une autre pellicule de silicium monocristallin d'une épaisseur de $200 \mu\text{m}$. Ce substrat est assez robuste

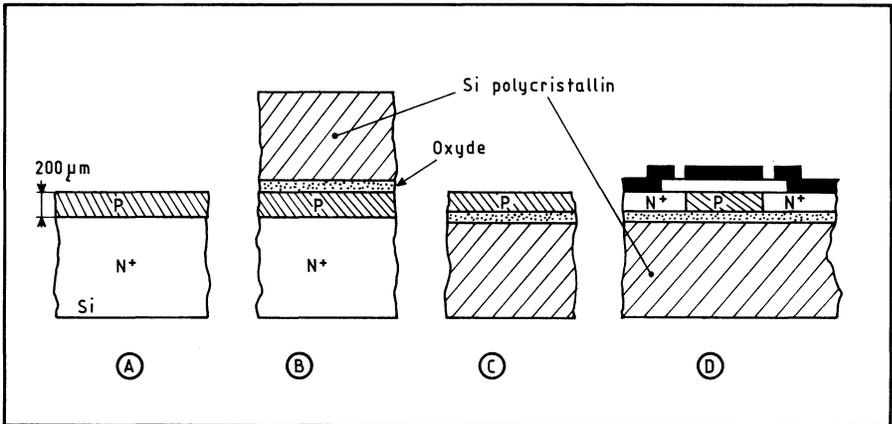


Fig. 26. — L'une des méthodes employées pour produire des transistors MOS dans les couches minces du silicium. Au lieu de silicium des types N^+ et P , on peut aussi utiliser du silicium des types P^+ et N .

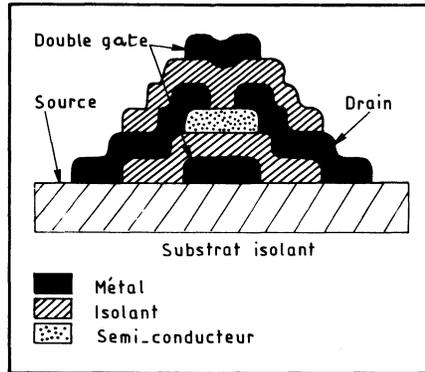
pour tenir convenablement pendant les différentes opérations qu'il faut encore entreprendre, en particulier l'attaque électrochimique et la diffusion à température élevée.

Le film de la fabrication est le suivant (on pourrait inverser les types de silicium). En (a) on fait croître une couche mince épitaxiale de silicium type P sur une plaquette monocristalline de silicium N^+ (épaisseur $200 \mu\text{m}$). En (b), on oxyde la couche épitaxiale et l'on fait évaporer sur elle du silicium polycristallin. La pellicule d'oxyde et le silicium polycristallin (hachures croisées) forment ensemble le substrat d'isolation. Puis, en (c), la couche N^+ est éliminée par attaque électrochimique; cette attaque est sélective : elle cesse aux limites de la région épitaxiale. Une morsure chimique complète cette opération (remarquer l'inversion haut pour bas de la figure). Enfin, en (d), les transistors MOS sont produits sur la couche de silicium par la méthode habituelle de masquage et des techniques de diffusion.

Les TFT du CNET

Des transistors à film mince (souvent appelés TFT à partir des initiales anglo-saxonnes de « Thin Film transistor ») ont aussi été développés par le CNET, à Lannion, afin de commander des affichages ⁽¹⁾. Ils ont été présentés au *Salon International des Composants*, de Paris, en 1978.

Fig. 27. — Structure du transistor au sélénure de cadmium, à l'alumine et au molybdène, développé au CNET-Lannion.



Les études ont été menées à la fois pour approfondir la physique du composant et réaliser des circuits pratiques. La structure qui a été adoptée pour le TFT est donnée figure 27. Le CNET a utilisé pour cela un évaporateur standard possédant un canon à électrons multicreusets associé à un changeur de masques sous vide. Le transistor est réalisé au cours d'un seul cycle de pompage. Les matériaux constitutifs sont le sélénure de cadmium, l'alumine et le molybdène respectivement comme semi-conducteur, isolant et métal. Le substrat est en verre. Les caractéristiques électriques du transistor dépendent non seulement des dimensions

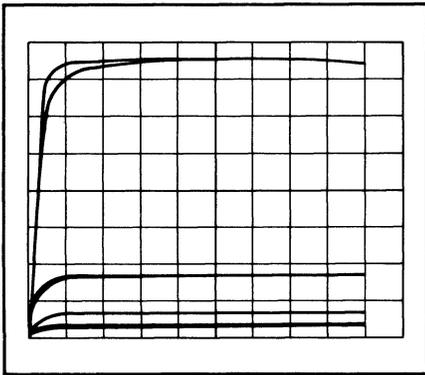


Fig. 28. — Exemple de caractéristiques du transistor.

géométriques imposées par le masque (longueur et largeur du canal), mais aussi de l'épaisseur du CdSe et de Al_2O_3 ainsi que de la structure cristalline du semi-conducteur, de sa stœchiométrie ainsi que de celle de l'isolant. Ces dernières propriétés des matériaux sont déterminées par les conditions d'évaporation et de recuit. On peut voir un exemple de caractéristiques sur la figure 28.

1. Voir : « Des transistors en couches minces pour les écrans de visualisation », par F. MORIN, du CNET, dans : « Électronique et Applications Industrielles » n° 254, du 1-6-1978.

Des caractéristiques électriques du TFT, on peut déduire la mobilité des porteurs dans le film semiconducteur. On trouve typiquement une mobilité électronique de l'ordre de $70 \text{ cm}^2/(\text{V.s.})$, alors que dans le canal d'un MOS celle-ci peut dépasser $500 \text{ cm}^2/(\text{V.s.})$. On voit tout de suite que cette plus faible mobilité va conduire à un temps de transit dans le canal plus élevé que dans un MOS, et on peut s'attendre à une fréquence de coupure plus basse. Par contre, la hauteur des barrières intercrystallines étant modulée par le potentiel de gate, on conçoit que la conductivité du canal va varier dans des proportions plus importantes que dans un MOS. On obtient aisément des rapports $R_{\text{OFF}}/R_{\text{ON}}$ supérieurs à 1 000. Mais le véritable avantage des TFT est qu'ils peuvent être réalisés sur de grands substrats peu chers.

En effet, étant donné que la fabrication ne requiert pas de températures élevées (la température de recuit peut ne pas dépasser $350 \text{ }^\circ\text{C}$), un grand nombre de matériaux peuvent être satisfaisants, pour peu qu'ils soient isolants et possèdent un bon état de surface (les verres et certains plastiques). On comprend alors pourquoi les applications principales du TFT sont surtout dans le domaine de la visualisation par écran plats.

L'écran prototype réalisé au CNET-Lannion est constitué d'une matrice de transistors (100 lignes, 10 colonnes, un TFT à chaque point de croisement). Chaque transistor agit comme un commutateur vis-à-vis d'un cristal liquide fonc-

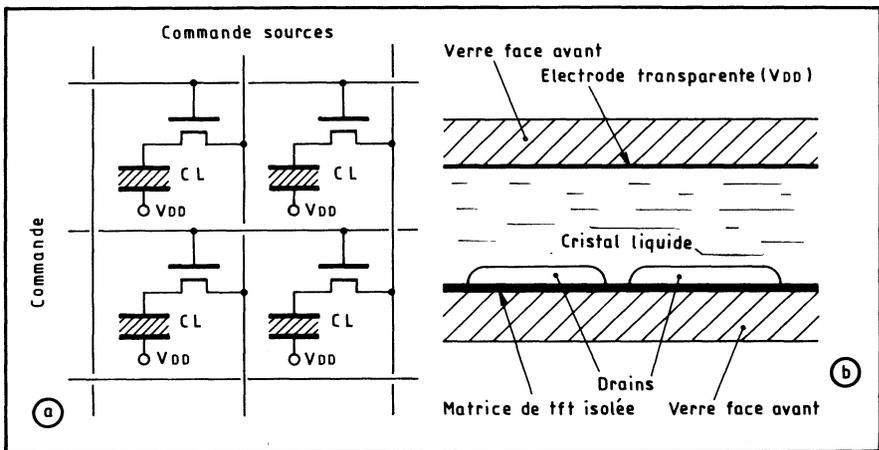


Fig. 29. — Schéma de la matrice développée au CNET pour l'affichage.

tionnant en régime de diffusion dynamique. Dans cette application, ce sont les propriétés d'amplificateur de tension et de grande résistance d'isolement qui sont utilisées dans ces transistors. On peut concevoir cependant d'autres fonctions (mémoire par exemple). On voit sur les figures 29 a et b les schémas de la matrice. La matrice développée comporte 1 000 transistors et 4 points/mm². Par la suite, il a été envisagé de remplacer le cristal liquide par une couche électroluminescente. Outre la matrice active, le CNET étudiait les registres à décalage qui devront y être associés. Un registre à 100 sorties permettrait d'adresser ainsi le signal vidéo en chaque point de la matrice.

Un autre domaine exploré est celui des hautes tensions. En effet, dans beaucoup d'applications, on désire intégrer des commutateurs haute tension (plusieurs centaines de volts). Cela semble possible avec les TFT.

14. LE HMOS

Pour accroître aussi bien la densité d'intégration que les performances, *Intel* a adopté en 1977 le concept HMOS, pour « *MOS à Hautes Performances* ». Selon ce concept, toutes les dimensions des transistors intégrés sont réduites selon la même échelle, ce qui permet de concevoir des MOS plus petits, plus rapides, et consommant moins. Par exemple, la longueur de canal n'aura plus que $4 \mu\text{m}$, l'épaisseur

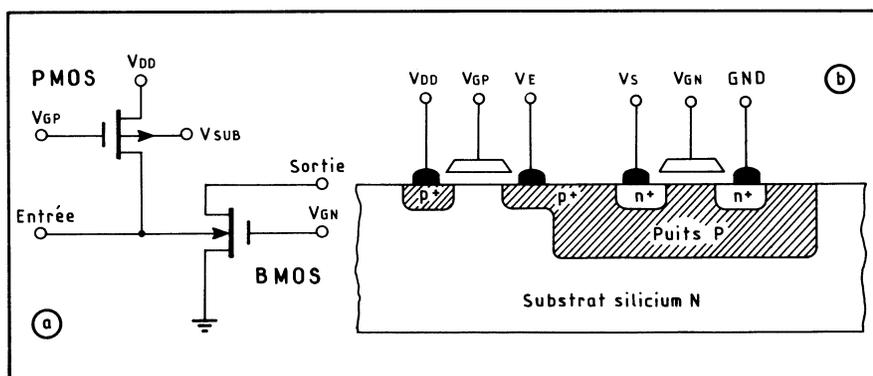


Fig. 30. — Structure et coupe de la porte BMOS

d'oxyde de gate sera ramenée à 600 ou 700 angstroms, les profondeurs des jonctions seront inférieures à $1 \mu\text{m}$, le reste à l'avenant; cette réduction à l'échelle 1/3 environ triple les performances : le facteur de qualité passe à 1 picojoule contre 3 pJ pour les MOS à canal de $6 \mu\text{m}$.

Ce sont les mémoires qui, les premières, ont bénéficié de cette conception, avec la 2115 A, RAM statique à temps d'accès de 25 ns. Ce processus, selon *Intel*, pourrait mener à des facteurs de qualité de l'ordre de 0,15 picojoule et à des RAM de 65 K bits sinon davantage.

Pour réaliser des HMOS interviennent uniquement des considérations de fabrication selon des techniques fines. Diverses sociétés ont suivi dans cette voie, et avec *Intel*, comptent aller plus loin dans la réduction des dimensions. Peut-être faudra-t-il alors réduire également les tensions d'alimentation et passer, par exemple, à +3 V là où aujourd'hui on s'en tient à 5 V, au risque de perdre tous les avantages de la compatibilité TTL.

S'opposant à la HMOS, la technologie VMOS sera présentée dans un chapitre spécial. Pour des performances semblables, elles n'obligeraient pas à une réduction des tensions d'alimentation, et par conséquent des seuils, mais on lui reprochait encore en 1978 d'être inexplorée et de ne pas avoir fait ses preuves. L'avenir appartiendra peut-être à des formes mixtes, puisqu'il n'est nullement exclu d'associer aux VMOS la philosophie de réduction d'échelle des HMOS.

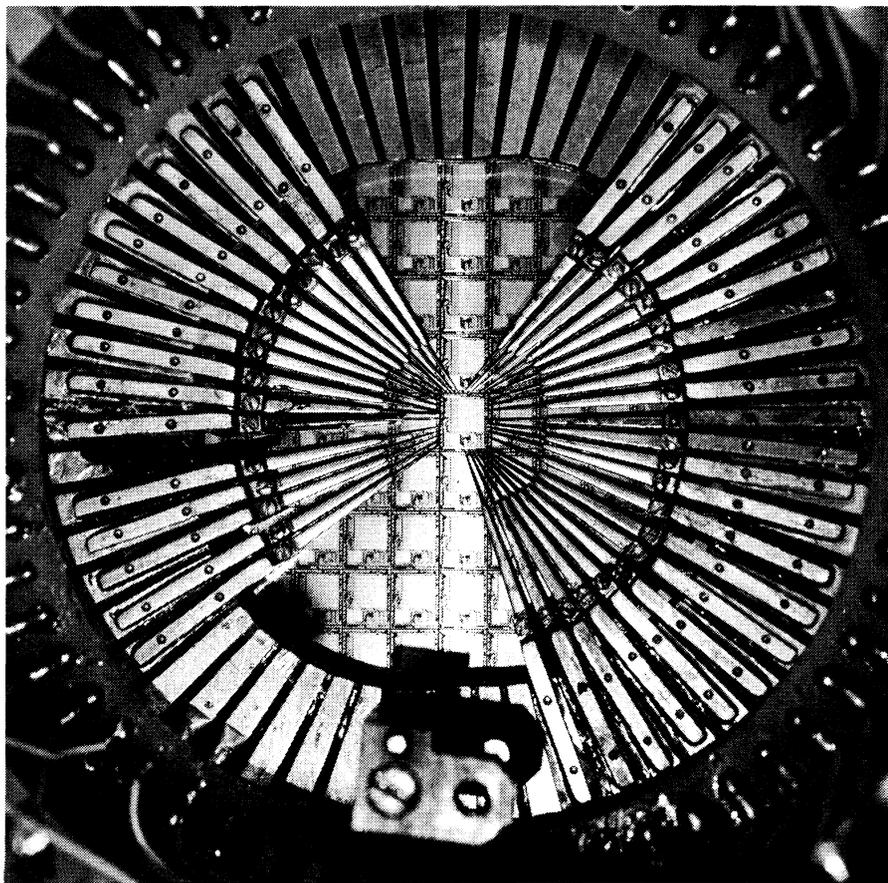
Toutes deux, HMOS et VMOS mènent ainsi à de grandes densités d'intégration, de grandes vitesses de travail comparables à celles des bipolaires et se contentent d'une seule tension d'alimentation.

15. LA STRUCTURE BMOS

Pour bien montrer l'abondance des structures possibles, on va encore en évoquer une, la BMOS. Imaginée par *Hitachi*, la structure logique BMOS, pour « *back-gated MOS* » permet d'accepter des niveaux de tension très faibles à l'entrée, 0,1 V par exemple, et par conséquent bien inférieurs à la tension de seuil courante des MOS. La porte de base se compose (fig. 30) d'un driver BMOS et d'un PMOS classique servant de source de courant. La coupe d'une cellule (en *b*) montre qu'on obtient une intégration de plus grande densité qu'avec les CMOS, ce au point qu'*Hitachi* espérait réaliser ainsi des mémoires RAM de 64 K bits. Le temps de propagation de la porte serait inférieur à 50 ns pour une consommation de l'ordre de 0,5 μ W, soit un facteur de qualité de 0,01 pJ annoncé en 1977.

16. LES BIMOS

Les structures BIMOS, enfin (à ne pas confondre avec BMOS), résultent du mariage sur une même puce de Bipolaires et de MOS. Ils ont mené tout d'abord à des amplificateurs opérationnels de très hautes performances qui seront présentés dans le chapitre consacré aux applications linéaires, tout comme les BIFET d'ailleurs (qui résultent, eux, de l'association de BIpolaires et de FET à jonction).



Test sous pointes des circuits intégrés, sur 40 contacts simultanément. (Doc. ITT-Intermetall.)

CHAPITRE VI

LES JFET ET LEURS FAMILLES

Les JFET sont des FET à jonction ; leur électrode de commande est en contact direct avec le semiconducteur. Ils interviennent surtout dans les commutateurs et les amplificateurs opérationnels (de technologies mixtes), et pourraient jouer un rôle important en versions MESFET ou VFET (Gridistor) dans les hautes et très hautes fréquences, ou même en puissance (VFET).

1. LE JFET DE BASE

Le JFET est également un transistor à effet de champ, mais à la différence des MOS, son électrode de commande est directement plaquée sur le semiconducteur et n'agit plus à travers une couche d'isolant. On verra que l'entrée est alors équivalente à une diode polarisée en inverse.

Sous forme de composant discret, le JFET se présente en batonnet ou en plaquette (fig. 1 *a*). Sa version intégrée est donnée en *b* et son symbole général, en *c*. On a proposé, ici, un JFET à canal N. Entre ses deux électrodes, notées *source* et *drain*, les porteurs majoritaires (qui sont ici les électrons des zones N) empruntent un chemin conducteur, le *canal*, et commandé par l'électrode de commande G, le *gate*.

Le gate étant relié à la source (donc $V_{GS} = 0$), on trace la caractéristique typique du JFET donnant le courant drain (I_D) en fonction de la tension drain (V_{DS}). On y distingue (fig. 2) une zone ohmique, linéaire, une région dite « de pincement » où le courant reste sensiblement constant, et enfin le claquage par avalanche. En agissant sur V_{GS} , on provoque une translation des courbes (fig. 3). La polarisation *négative*, ici, appliquée au gate, montre que l'on attaque en inverse la diode d'entrée. Le canal du JFET est comparable à un conducteur dont la section diminue quand on augmente le champ électrique de commande, d'où le nom de transistor à effet de champ.

Le courant drain I_D est lié à V_{GS} par :

$$I_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right)^2$$

avec :

I_{DSS} = courant drain pour $V_{GS} = 0$, encore appelé *courant de saturation*;
 V_P = tension de pincement, qui détermine ce que l'on appelle *l'état bloqué* : c'est la tension V_{GS} qui supprime le courant I_D (et qui serait de quelque $-5V$ avec les courbes données).

Il est intéressant de tracer la caractéristique $I_D = f(V_{GS})$, comme avec les tubes (fig. 4). La transconductance, ou *pente*, notée g_m , se définit par le rapport $\Delta I_D / \Delta V_{GS}$ établi, par exemple, entre les points C' et D' de la courbe, soit :

$$g_m = \frac{\Delta I_D}{\Delta V_{GS}}$$

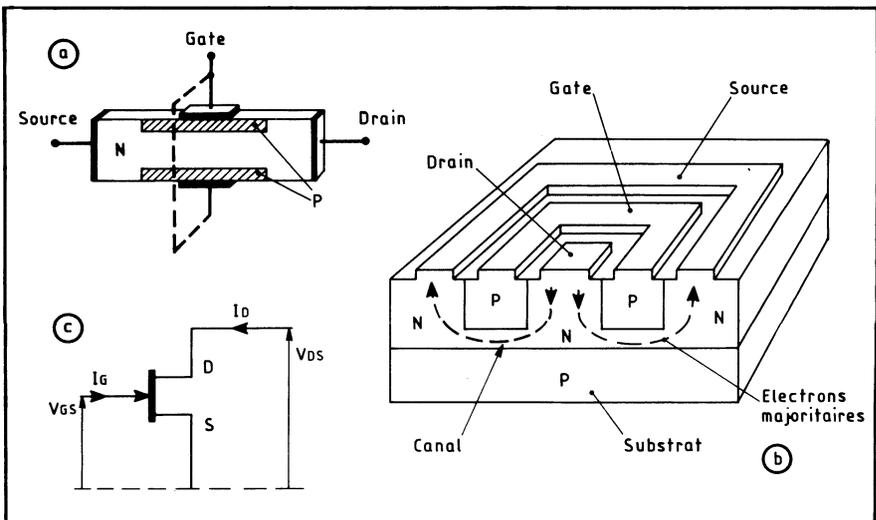
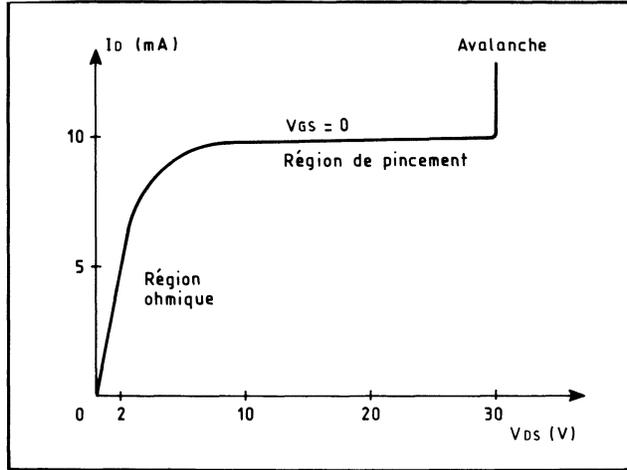


Fig. 1. — FET à jonction en plaquette (a), en intégré (b) et symbole du FET à canal N (c).

Fig. 2. — Courbe type du JFET donnant le courant drain en fonction de la tension drain-source.



ce qui donne, si l'on développe à l'aide de la formule précédente :

$$g_m = 2 I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right) = g_{m0} \left(1 - \frac{V_{GS}}{V_P} \right)$$

$$= g_{m0} \sqrt{\frac{I_D}{I_{DSS}}}$$

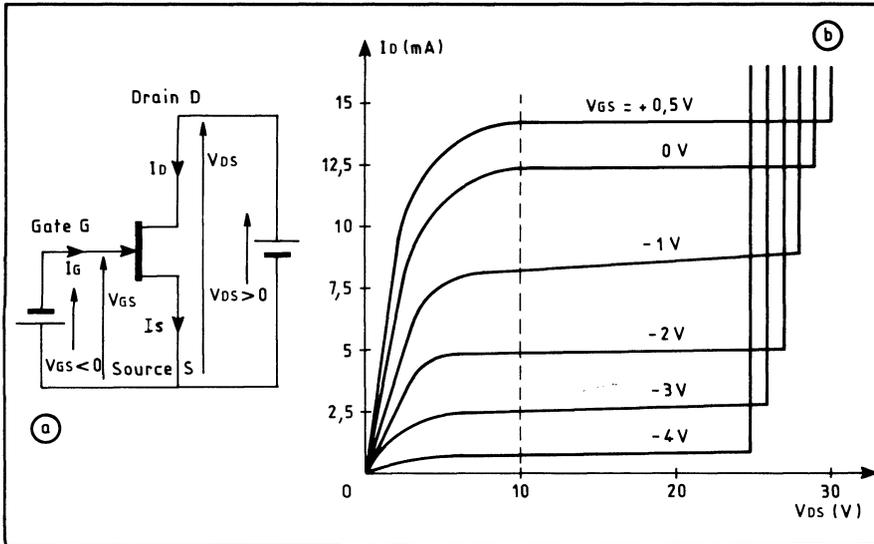


Fig. 3. — Réseau de courbes Idrain = f (tension drain-source) pour plusieurs tensions de gate.

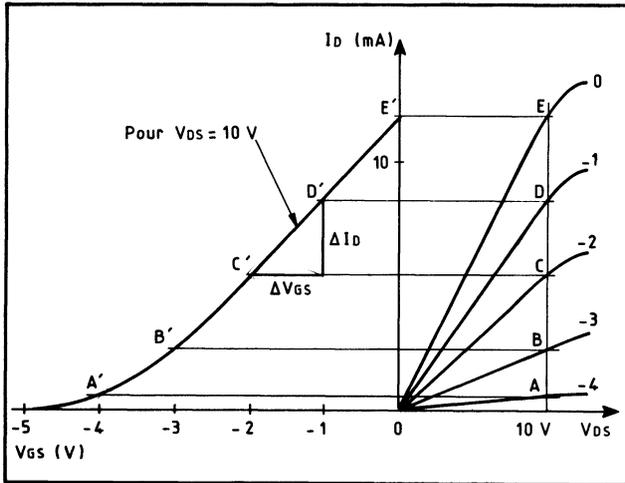


Fig. 4. — Courant drain en fonction de la tension gate-source. Sa pente est la transconductance.

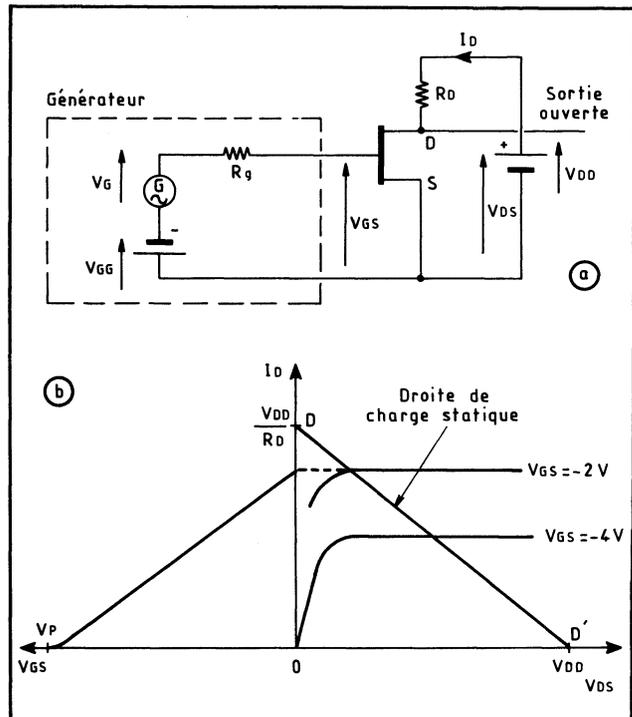


Fig. 5. — Tracé de la droite de charge statique.

Le courant I_D étant égal à I_{DSS} pour $V_{GS} = 0$, on a alors :

$$g_m = g_{m0} = \frac{2 I_{DSS}}{-V_p}$$

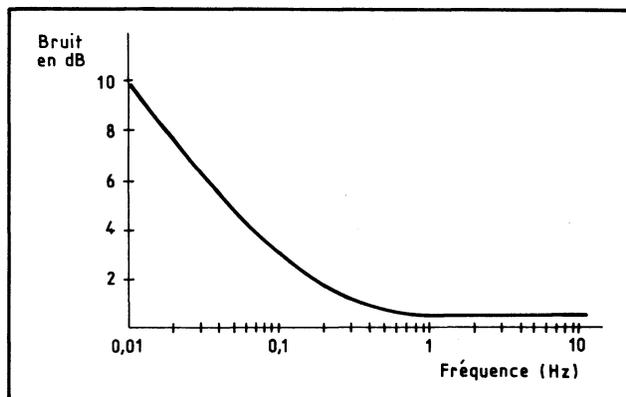
Ce qui donne la transconductance maximale, notée g_{m0} (V_p étant négatif, le signe — fournit un résultat positif).

La droite de charge est obtenue d'après le montage de la figure 5 a, en b.

Dans le cas d'un JFET à canal P, les courbes sont identiques mais inversées, bien entendu, eu égard aux tensions appliquées (négative au drain et positive au gate).

Les variations de température provoquent généralement une réduction du courant drain pour une température croissante, de l'ordre de 0,6 %/°C pour les JFET discrets les plus usuels. Ce coefficient de température *négatif* les protège efficacement de l'emballage thermique.

Fig. 6. — Courbe de bruit typique du JFET.



Les caractéristiques de bruit des JFET sont généralement très favorables; ce bruit résulte de l'agitation thermique dans le canal conducteur, auquel s'ajoute le bruit de grenaille (« shot noise ») créé par le courant de fuite de gate I_{GSS} . Le bruit en $1/f$ est limité aux alentours de 100 Hz généralement (fig. 6). Les JFET sont ainsi moins « bruyants » que les MOS.

2. LE JFET SCHOTTKY

Le FET Schottky est un JFET dont l'électrode de commande est constituée par une diode Schottky, c'est-à-dire une jonction métal-semiconducteur faiblement dopé (fig. 7).

Le fonctionnement du FET Schottky est tout à fait semblable à celui des transistors à effet de champ à jonction décrits par SHOCKLEY : une tension négative

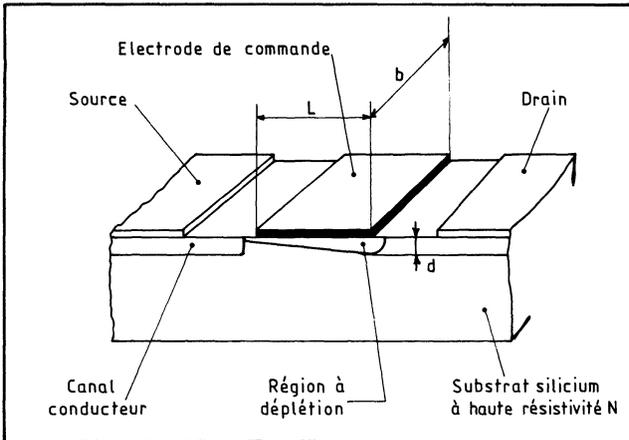


Fig. 7. — Le transistor MESFET, ou transistor à effet de champ à attaque Schottky; l'électrode de commande métallique est directement appliquée sur le silicium.

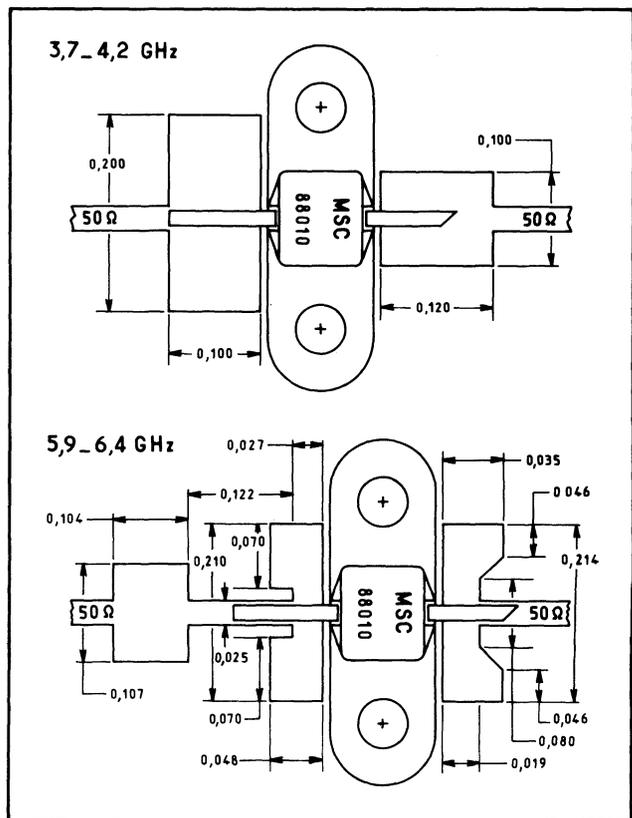


Fig. 8. — Montage du MESFET à L'AsGa, type MSC 88010, en hyperfréquences. Attention : toutes les dimensions sont en pouces, avec 1 pouce = 25,4 mm.

appliquée au gate, l'électrode de commande, accroît l'épaisseur de la zone de déplétion (zone désertée) par laquelle passe le courant. Dès que la tension de drain devient suffisante pour que la quasi-totalité de la zone soit désertée, il y a pincement et le courant ne croît plus avec la tension.

Proposé en 1966 par C. A. MEAD, le FET à jonction Schottky semble avoir été initialement réalisé par K. E. DRANGEID et son équipe d'IBM, à Zurich, sous le nom de MESFET (« métal, silicium, FET »). Son intérêt réside dans sa fréquence élevée de travail, 5 GHz dès l'origine avec des MESFET sur silicium, davantage avec des éléments à l'arséniure de gallium (AsGa, ou GaAs selon la terminologie américaine).

Ces FET Schottky à l'AsGa viennent de connaître un développement spectaculaire en hyperfréquences, notamment dans la voie des transistors à faible bruit (vers le décibel à 10 GHz) et des fortes puissances. Au début de 1978, divers laboratoires travaillaient sur des transistors 10 GHz de plusieurs watts. Un élément 2,5 W à 8 GHz étant même commercialisé (le MSC 88010 de *Microwave Semiconductor Corp.*).

Les FET Schottky interviennent comme oscillateurs, mélangeurs, modulateurs... dans des applications qui relèvent des télécommunications par satellites (en attendant la réception directe TV), le radar, les contre-mesures électroniques pour les militaires; peut-être trouveront-ils également leur place dans les systèmes numériques très rapides.

Les FET hyperfréquences sont naturellement des composants à large bande. Celle-ci dépend essentiellement des géométries de leurs circuits externes. Par exemple, le MSC 88010 fonctionnera de 3,7 à 4,2 GHz dans le montage de la figure 8 *a* (où toutes les dimensions sont données en pouces, avec 1 pouce = 25,4 mm), et de 5,9 à 6,4 GHz avec celui de la figure *b*. Travaillant en classe A avec 1 dB de compression, son gain atteint 6 dB à 8 GHz et sa bande passante propre va de 2 à 10 GHz. Son boîtier est du type métal-céramique, à faibles pertes, et sa résistance thermique a pu être abaissée à 10 °C/W. Sous une tension drain-source de 5 V, il consommera typiquement 1,5 A et sa transconductance sera de 400 milimhos.

En versions faible bruit, un transistor commercialisé tel que le HFET-1101 de *Hewlett-Packard* offrirait un bruit de 1,6 dB à 4 GHz pour un gain de 16 dB à la même fréquence et une puissance de 15,5 dBm.

L'Astec de Thomson-CSF

La fabrication du FET Schottky, telle que proposée par les Laboratoires de recherches de Corbeville, de *Thomson-CSF*, comprend les étapes principales suivantes (fig. 9) :

- une couche d'oxyde épais est gravée; on crée ensuite une couche d'oxyde mince;
- par implantation ionique à travers l'oxyde mince, on fabrique une zone N;
- un masquage à l'or limite une seconde implantation ionique N+;
- on retire le masquage à l'or, on renforce les couches d'oxyde et l'on ouvre à nouveau des fenêtres;
- les interconnexions sont obtenues par une métallisation molybdène-or.

On réalise ainsi un « Astec » (pour « TEC à Accès Schottky ») qui, chargé par une simple résistance, constitue une porte inverseuse élémentaire (fig. 10); cette résistance est également du type FET.

La tension de seuil V_{TH} du FET dépend de ses caractéristiques et, en particulier, de la hauteur de la barrière métal-silicium; elle est ainsi de l'ordre de 0,6 à 0,7 V.

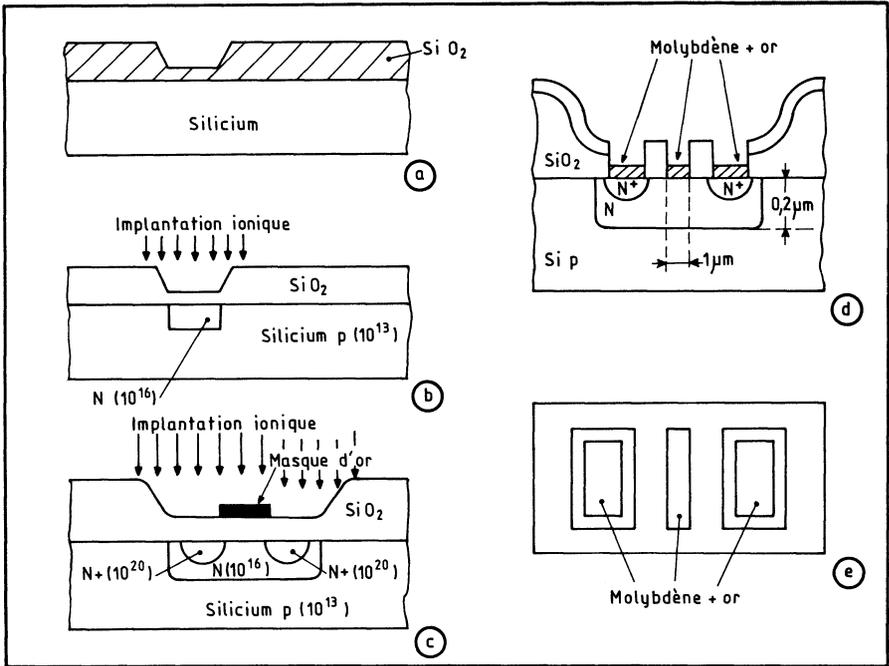


Fig. 9. — Quelques étapes dans la fabrication d'un élément Astec.

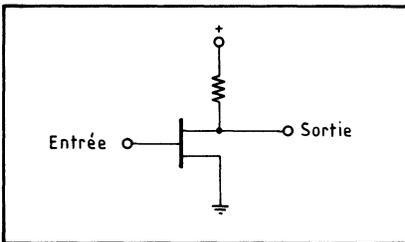
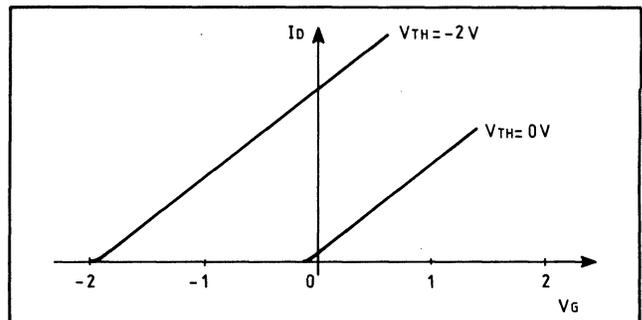


Fig. 10. — Principe de la porte à FET Schottky.

Fig. 11. — Selon la tension de seuil du FET, le transistor MEFSET peut être ou non bloqué.



Or, la diode Schottky qui travaille en inverse doit, pour rester bloquée, être polarisée en dessous de 0,4 V. Il s'ensuit deux modes de fonctionnement possibles :

- 1° Avec $V_{TH} = 2 \text{ V}$ et tension d'alimentation $V_A = 6 \text{ V}$;
- 2° Avec $V_{TH} = 0 \text{ V}$ et $V_A = 1 \text{ V}$. Mais dans ce cas, la diode Schottky reste en direct.

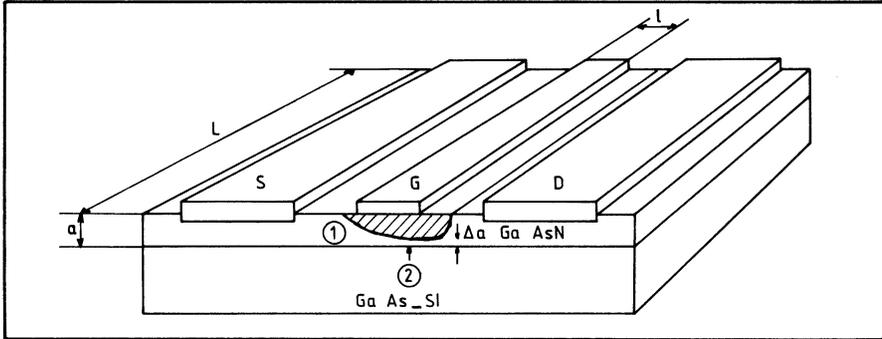


Fig. 12. — FET Schottky à substrat d'arséniure de gallium.

Ces deux cas se traduisent par les courbes donnant le courant drain en fonction de la tension de commande V_G (fig. 11).

La haute densité permise par cette technologie et sa faible consommation permettent d'imaginer les fonctions et systèmes intégrés les plus complexes. En effet, une porte élémentaire n'occuperait que $10 \times 10 \mu\text{m}$, et une cellule-mémoire $30 \times 30 \mu\text{m}$. Ainsi, 1 mm^2 de silicium suffirait pour 10 000 portes ou 1 000 bits-mémoire, et une puce de 25 mm^2 pour plus de 250 000 portes ou 25 000 bits-mémoires, valeurs prodigieuses mais non surprenantes pour qui a suivi l'évolution de l'intégration.

D'ailleurs, fonctionnant en numérique, la porte fondamentale peut aussi bien travailler en linéaire et pourrait servir à réaliser des amplificateurs opérationnels, du continu aux VHF (1 GHz, par exemple). L'utilisation de FET à pente élevée (10 mA/V) permettrait d'augmenter la puissance.

Passant à l'arséniure de gallium, Thomson-CSF développait, en 1977, avec le soutien de l'Administration, des FET, moyen et fort niveau, pour utilisation en amplification et oscillation à partir de 5 GHz. Ces dispositifs comprenaient alors le TH 467 qui délivre une puissance de sortie de 500 mW à 5 GHz avec un gain maximum supérieur à 6 dB; le TH 469 qui délivre une puissance de sortie de 300 mW à 10 GHz avec un gain maximum supérieur à 5,5 dB; ainsi que le TH 470 qui délivre une puissance de sortie de 100 mW à 13 GHz avec un gain maximum supérieur à 4,5 dB; cela dans une série qui devait être développée.

Le TBM du LEP

Le LEP a également étudié des structures à l'AsGa, appelées TBM (pour transistor à barrière métallique) dont des versions très performantes ont ensuite été commercialisées par RTC. La structure du TBM est représentée figure 12, le

tableau A montre qu'avec l'AsGa, en fonction de la concentration en électrons libres Nd :

- la mobilité μ des électrons est double de celle du germanium, et quintuple de celle du silicium;
- la saturation de vitesse des porteurs V_{sat} en régime de fort champ (qui limite le courant maximal) est supérieure;
- la tension de claquage V_B est plus élevée.

Tableau A. — Quelques caractéristiques du silicium, du germanium et de l'arséniure de gallium

MATÉRIAU	POUR CONCENTRATION EN	MOBILITÉ μ	SATURATION DE	TENSION DE
	ÉLECTRONS LIBRES			
	Nd de	cm ² /(v.s.)	V_{sat}	V_B
	(at/cm ³)		(cm/s)	(V)
Ge	10^{16}	3200	$0,65 \cdot 10^7$	25
	10^{17}	2000		5
Si	10^{16}	1200	$0,9 \cdot 10^7$	55
	10^{17}	700		12
AsGa	10^{16}	6000	$1,2 \cdot 10^7$	70
	10^{17}	4800		16

En effet, la théorie classique de SHOCKLEY montre que la fréquence de travail du FET est proportionnelle au quotient de la mobilité des porteurs par le carré de la longueur du gate, si celle-ci est égale ou supérieure à 4 μm . En dessous, et c'est le cas des FET Schottky puisque les dimensions sont couramment de 1 à 2 μm , le champ dans la zone du canal est tel que les porteurs atteignent leur vitesse de saturation; en conséquence, celle-ci doit être le plus élevé possible.

La réalisation du TBM, selon l'analyse qu'en ont fourni Pierre BAUDET et Michel BINET ⁽¹⁾ est la suivante. Le matériau de base est un monocristal d'arséniure de gallium semi-isolant dopé au chrome ($\rho > 10^8 \Omega \cdot \text{cm}$) et poli par voie mécano-chimique. La croissance de la couche épitaxiale est effectuée en phase vapeur dans un tube de quartz de très haute pureté afin d'introduire un minimum d'impuretés compensatrices. La couche de type N a une concentration comprise entre 3 et $10 \cdot 10^{16}$ atomes par centimètre carré pour une épaisseur inférieure à 0,5 μm .

L'isolation des transistors est obtenue par décapage chimique jusqu'au substrat (technique « mesa »). Les contacts ohmiques de « source » et de « drain » sont réalisés par alliage d'or-germanium (88 %-12 % en poids) recouverts d'une fine

1. Du LEP, dans un article publié par *L'Onde Électrique*, 1974, vol. 54, n° 1, pp. 31-35.

couche d'or (1 000 Å). La diode Schottky de commande est réalisée en aluminium de 0,5 µm d'épaisseur évaporé sous ultra-vide; afin de réduire au minimum la capacité d'entrée, les connexions de gate sont rejetées sur le substrat semi-isolant.

Afin d'éliminer les éléments parasites, la « puce » est montée directement sans boîtier dans un système microstrip. Ce système se compose d'un bloc de cuivre doré surmonté d'un rail, de part et d'autre duquel sont fixées deux lignes 50 Ω sur alumine; la « puce » est soudée sur le rail entre les deux lignes. Les connexions sont réalisées en fil d'or de 25 microns de diamètre soudé par thermocompression. En l'absence de transistor, l'isolation entrée-sortie est supérieure à 40 dB en bande X.

Des performances-types obtenues sont :

— fréquence de coupure	: 25 GHz	
— pente intrinsèque	: 100 mA · V ⁻¹ · mm ⁻¹	
— facteur de bruit minimal	: 4,3 dB	} à 6 GHz
— gain associé au bruit minimal	: 8 dB	
— gain maximal	: 10 dB	
— bruit associé au gain maximal	: 6,7 dB	

Ces performances peuvent être améliorées de diverses façons : amincissement de la couche épitaxiale, augmentation du dopage, réduction des dimensions vers des électrodes de commande de 1 µm de largeur, etc.

3. LES GRIDISTORS ET LE VFET

Les versions unipolaires : le VFET

Bien qu'invention typiquement française, due au Docteur Stanislas Teszner, conseiller scientifique du *CNET* (*Centre National des Télécommunications*), le Gridistor (pour « transistor à grille ») dont une variante est le VFET (pour « FET vertical ») a commencé à être exploité industriellement, au Japon, à partir de 1973.

Le VFET n'est nullement comparable au VMOS, sauf en un point peut-être : le courant traverse le transistor en volume et non plus en surface. Ce transistor est caractérisé par une structure intégrée à canaux multiples verticaux et peut fournir une puissance très élevée par unité de surface, ou travailler à des fréquences dont la limite, entrevue en 1977-1978, se situerait vers les 40 GHz.

Les brevets sur les Gridistors ont été pris de 1956 en France (il s'agissait alors de son ancêtre, le Tecnetron) à 1973. Ce Gridistor est un transistor à effet de champ, à jonction, dans lequel l'électrode de commande est cette fois une « grille »; c'est ce que montrent les deux coupes de la figure 13 extraites des premiers brevets, la première pour un transistor conçu en technologie à double diffusion et la seconde, en technologie à couche épitaxiale. Comme tous les FET, il offre une grande impédance d'entrée; il présente également un faible niveau de bruit, n'introduit que très peu de distorsion, et résiste remarquablement aux surcharges. La figure 14 en montre une vue cavalière où, cette fois, la prise de contact de grille est placée dans un puits.

Dès 1973, les Japonais *Sony*, *NEC*, *Yamaha*... s'intéressaient au VFET pour des applications à la hi-fi. Ainsi, *Yamaha* réalisait en 1974 des VFET de puissance, soit de 10 W (tensions de 300 à 500 V, transconductance de 50 mA/V, coefficient

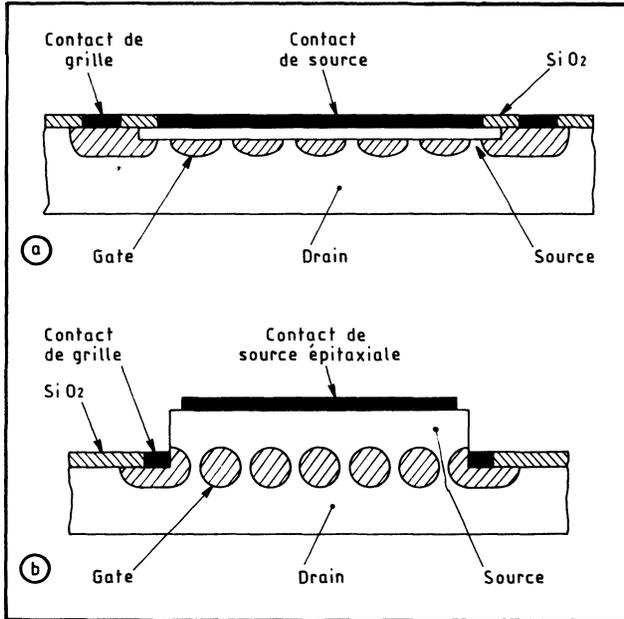


Fig. 13. — Les premiers Gridistors, obtenus par double diffusion (a) ou par épitaxie (b).

Fig. 14. — Vue cavalière d'un Gridistor, ou VFET, avec le contact de l'électrode de commande situé dans un puits.

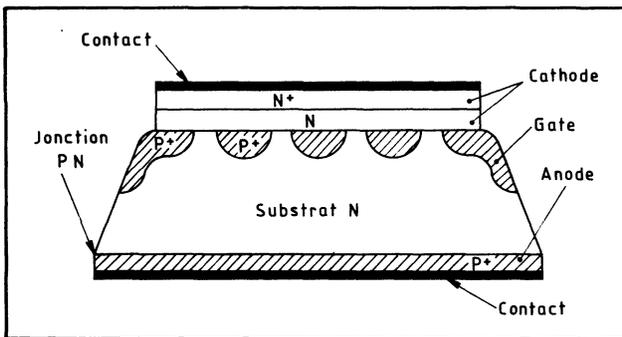
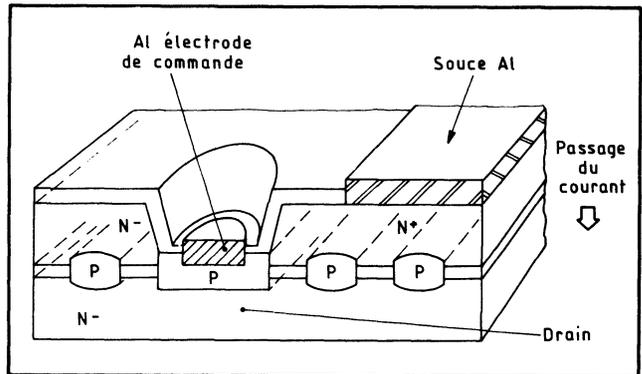


Fig. 15. — Structure du Gridistor bipolaire; le gate, ou grille, en constitue une électrode.

d'amplification de 50, résistance interne de $1\ 000\ \Omega$), soit de 300 W (10 A max, 200 à 300 V, coefficient d'amplification de 5).

En 1974, l'*US Air Force* proposait un VFET Schottky de 3 W à 28 GHz, avec traits de 1 micron. En France et dès la fin des années 60, *Sesco* puis *Sescosem*, sur contrats de recherche du *CNET*, de l'*US Navy* et de la *SEFT*, développaient le Gridistor et obtenaient en 1971 une puissance de sortie de 1 W à plus de 1 GHz, alors record mondial pour un FET. En 1973, *Sescosem* présentait au Salon des *Composants* de Paris le Gridistor TH454 fournissant 200 mW à 2,7 GHz et 500 mW à 2 GHz en classe A. Auparavant, *Jeumont* commercialisait (en 1963) un Tecnetron discret commandant 10 A sous 300 V inverses max.

L'implantation ionique, selon la technologie développée par le *CNET*, et les versions Schottky du VFET laissent prévoir des composants de ce type à produit puissance-bande passant dépassant 50 GHz avec une définition de trait de 0,5 microns (20 GHz étaient obtenus en 1978).

Versions bipolaires ou mixtes

Toutes les versions citées ci-dessus sont des variantes de FET, donc des transistors unipolaires. Or, le Gridistor se prête également à des structures bipolaires prévues dans un brevet dès 1961. C'est ce que montre la figure 15 où l'on se trouve en présence d'un PNP (auquel se superpose une structure de FET). La densité

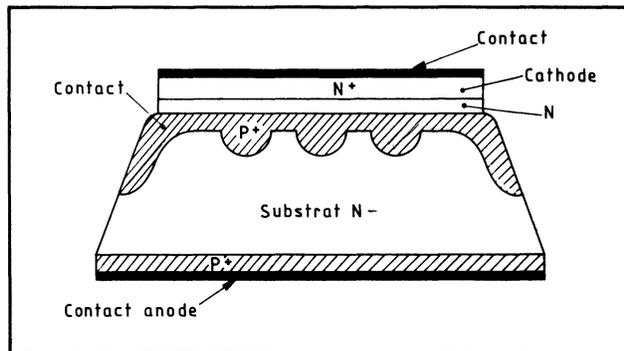


Fig. 16. — Variante bipolaire du Gridistor, fonctionnant en bistable : c'est l'équivalent d'un thyristor.

de courant peut atteindre $100\ \text{A}/\text{cm}^2$ et la coupure est assurée en bien moins d'une microseconde. La chute de tension directe est de 1,3 V et le transistor supporterait de 1 000 à 1 200 V. Il est désigné par le sigle de BIG, pour « bipolaire Gridistor » (son licencié était *Alsthom-Atlantique*).

En modifiant la structure de la grille de commande (fig. 16), on aboutit à un transistor à 4 couches, type thyristor, qui a été développé par *Jeumont-Schneider*.

Ce Gridistor constitue par conséquent une autre des filières des FET, riche de possibilités.

4. LES JFET EN COMMUTATION

En versions intégrées, les JFET ont tout d'abord été exploités en commutateurs, en particulier par *Siliconix*. Dans une telle application, le JFET est chargé de

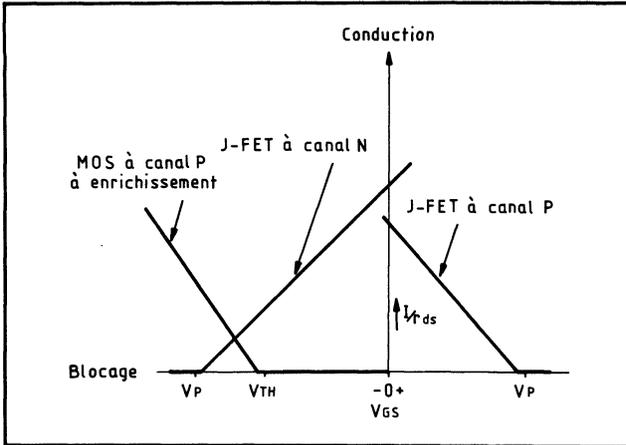


Fig. 17. — La conduction des FET en fonction de la tension de commande V_{GS} .

transmettre ou de bloquer des informations, analogiques ou numériques, lui-même faisant l'objet d'une commande par tout ou rien, donc numérique.

Reprenons le cas d'un JFET discret : une tension positive de commande rend conducteur le JFET à canal N et bloque celui à canal P, alors qu'une tension négative exerce l'effet inverse. Lorsque la tension gate-source V_{GS} est nulle, le JFET est conducteur (tout comme les MOS à déplétion).

Lorsque le FET est conducteur, le courant emprunte son canal qui lui oppose une résistance faible mais non négligeable; lorsque le FET est bloqué, cette résistance s'élève considérablement. La résistance drain-source r_{DS} dépend donc de la tension appliquée entre gate et source, V_{GS} , comme le montrent les courbes de principe de la figure 17. On remarque que :

- les J-FET, canal N ou P, sont conducteurs lorsque $V_{GS} = 0$;
- leur conduction décroît au fur et à mesure que V_{GS} devient négative pour le J-FET à canal N, positive pour celui à canal P;

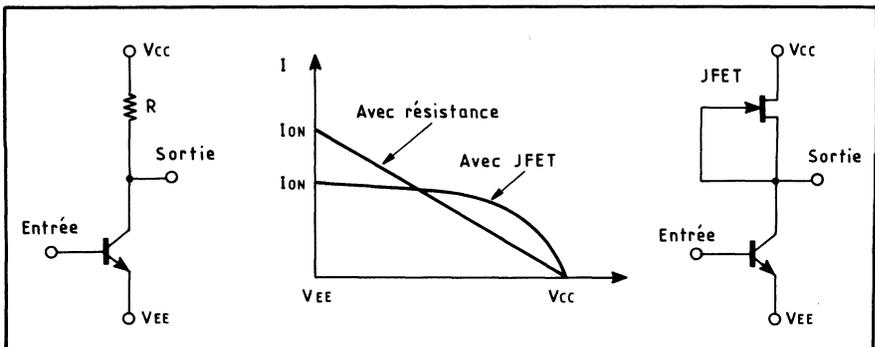
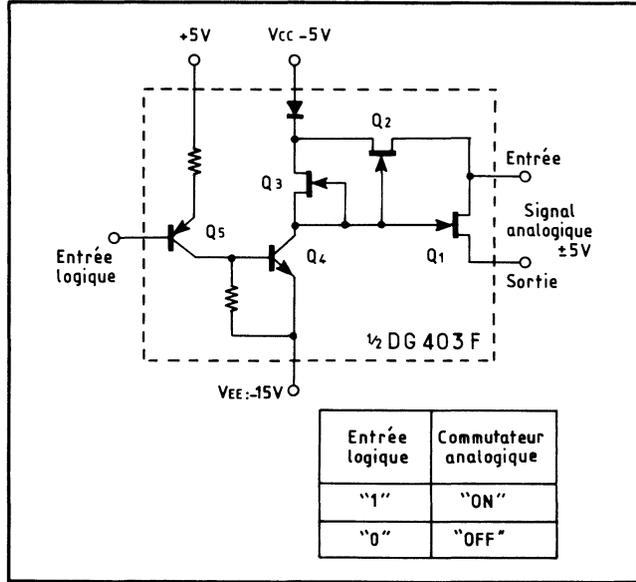


Fig. 18. — Le JFET en résistance de charge, comparé à une résistance ohmique.

— le blocage est obtenu pour une tension dite de *pincement*, V_p . Celle-ci est plus élevée, en valeur absolue, pour le J-FET à canal P que pour celui à canal N.

Par exemple, supposons qu'un signal analogique de +5 V est appliqué à la source et que le drain est à -3 V. Si $V_p = -8$ V, il faudra au moins -11 V appliqués au gate du J-FET à canal N pour le bloquer. Dans les mêmes conditions ($V_s = +5$ V et $V_D = -3$ V), mais avec $V_p = +8$ V, cette fois, il faudra au moins +13 V pour bloquer le J-FET à canal P.

Fig. 19. — Ensemble commutateur dissipant au minimum à l'état passant.



Le drain et la source d'un J-FET conducteur sont approximativement au même potentiel; la tension gate-canal de ce même élément est à peu près nulle. D'autre part et tout comme les MOS à déplétion, le JFET peut servir de charge dans des montages identiques. Ainsi, un transistor bipolaire peut être chargé par un JFET (fig. 18) et non plus par une résistance; on y gagne en consommation lorsque le bipolaire est conducteur (V_{EE} à la sortie), de l'ordre de 50 %, pour un temps de commutation inchangé.

A partir de ce dernier schéma, on peut comprendre le fonctionnement d'un commutateur de tensions analogiques, typique, tel que le DG403F dont le quart (c'est un élément quadruple) a été dessiné dans la figure 19. Le commutateur proprement dit est le JFET à canal N noté Q1; à l'état passant, son gate est court-circuité à la source tandis qu'à l'état bloqué, la tension V_{EE} lui est appliquée; les conditions sont : $V_p \leq -10$ V et V_a , tension du signal analogique : ± 5 V. A l'état passant de Q1, en effet, correspond Q4 bloqué. La chute de tension aux bornes de Q3 s'annule alors, ce qui permet à Q2 de se débloquer; en conséquence, V_{GS} de Q1 est nulle. Notons que la consommation est minimale pour le commutateur passant.

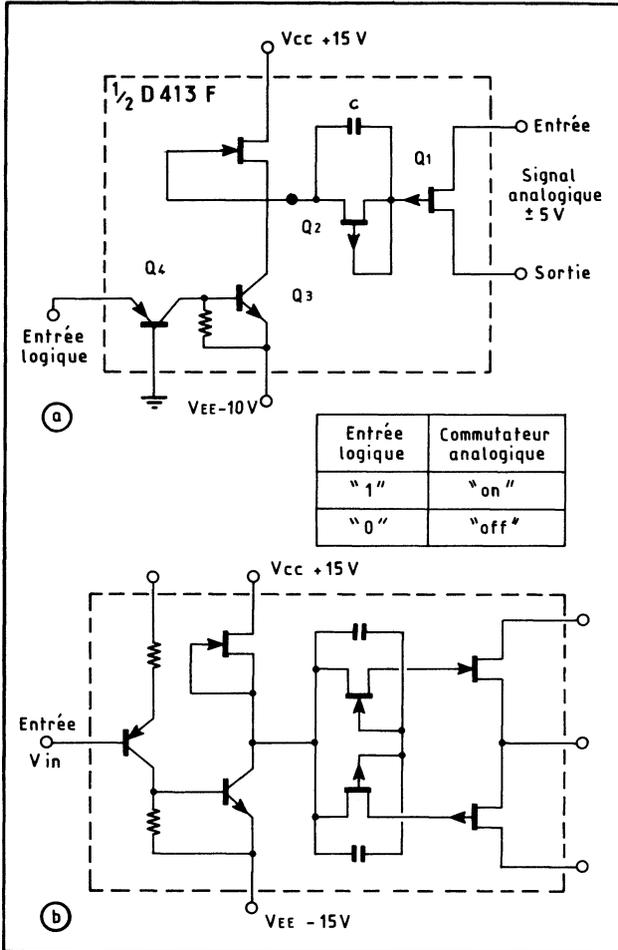


Fig. 20. — Grâce au FET limiteur Q_2 , le courant extrait du gate de Q_1 n'est que de $6 \mu\text{A}$ à l'état passant (a). Variante à montage symétrique (b).

On peut constater que la commande est bien numérique (niveaux logiques), alors que l'information commutée peut être, et sera souvent analogique. D'autre part, ce dispositif est mixte, en ce sens qu'il comprend des transistors bipolaires et unipolaires sur une même puce. C'est ce qui sera appelé technologie BIFET (« Bipolaire et FET ») dès que l'on passera aux amplificateurs opérationnels.

Afin de réduire le courant de commande d'un tel commutateur, on recourt aux schémas de la figure 20. En a, le FET, noté Q_2 , est inséré entre l'interface et le commutateur. Lorsque Q_3 est bloqué, le gate de Q_1 est à $+15 \text{ V}$ et Q_1 est bloqué. Lorsque Q_3 est conducteur, la tension V_{GS} de Q_1 devient négative et l'électrode de commande de Q_1 consomme un courant I_G que lui fournit sa source. Grâce au FET, Q_2 , monté en diode, ce courant de saturation est limité à $6 \mu\text{A}$. En b, on a réalisé un montage complémentaire à la sortie, piloté lui-même par des FET complémentaires agissant en limiteurs de courant.

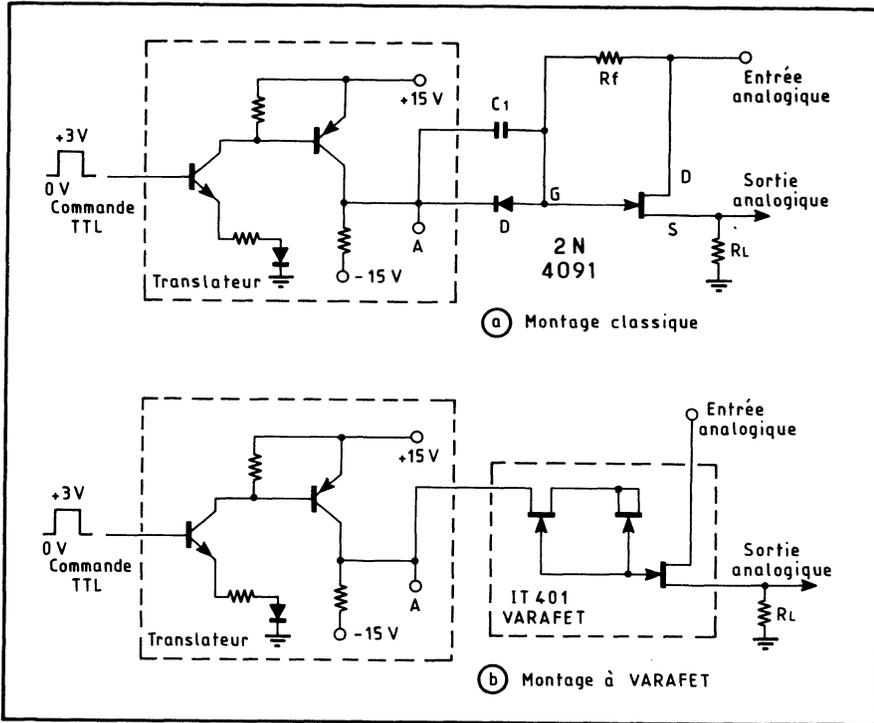


Fig. 21. — Structures comparées d'étages de commutation analogique utilisant un étage classique (a) ou un Varafet (b).

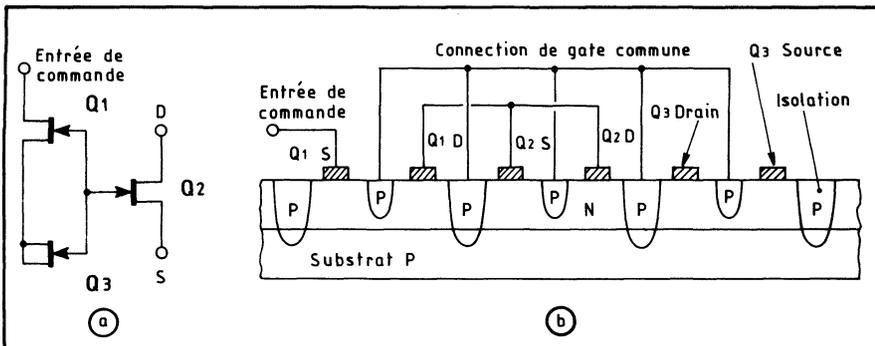


Fig. 22. — Organisation et structure du Varafet.

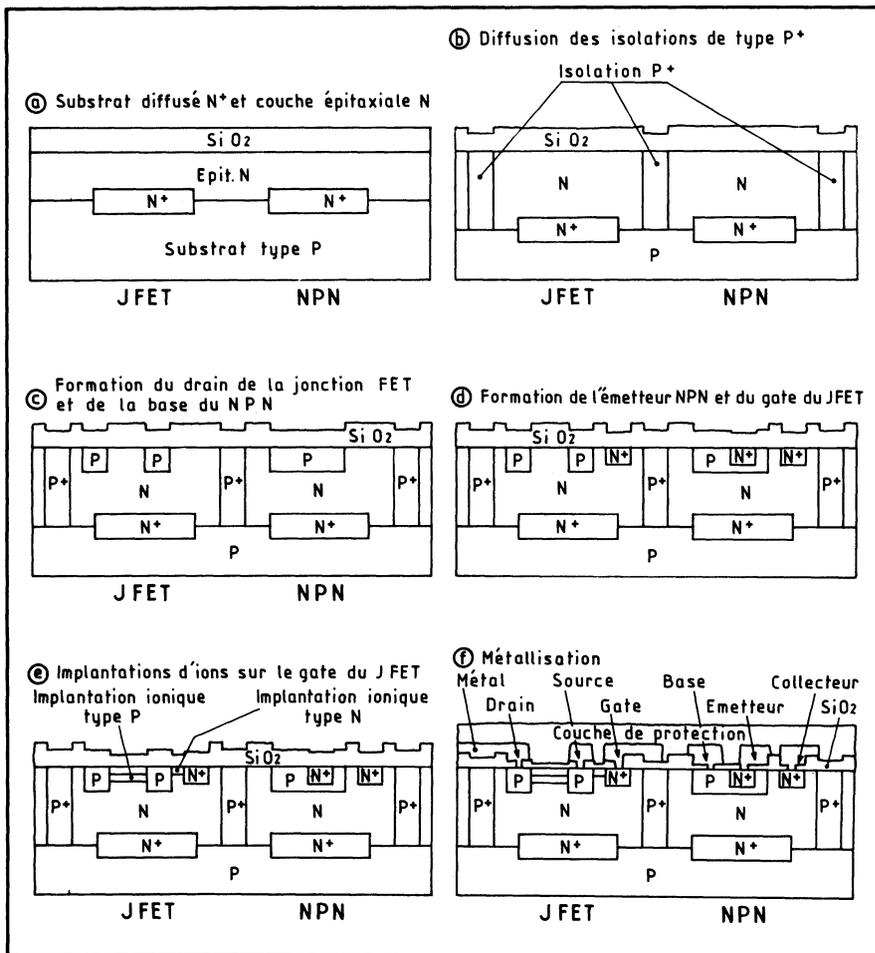


Fig. 23. — Fabrication d'un « BIFET ».

Le Varafet

Le Varafet ne constitue d'une variante des structures à FET, destinée essentiellement à servir de commutateur analogique. Il a été développé fin 1976 chez *Intersil* par Dick WILENKIN et commercialisé en 1978 ⁽¹⁾.

1. Une étude complète lui a été consacrée dans le n° 242, du 15-10-77, de la revue « *Électronique et Applications Industrielles* », p. 25 et la suite.

Circuit intégré élémentaire, il interviendra précédé par un translateur de niveau (TTL à JFET) comme le montre la figure 21 *b*, comparée au montage habituel *a*. On élimine ainsi le pont diviseur capacitif dû à la diode et aux jonctions gate-drain et gate-source de l'ancien montage (ce pont peut interdire au FET de devenir franchement conducteur, et ses valeurs varient avec le signal), et l'on supprime également le condensateur (qui limitait l'amplitude du signal analogique) et la résistance de rappel qui crée un autre pont.

Les deux FET d'entrée du Varafet jouent un rôle de diode de protection, le second (Q_2 , fig. 22) ayant été construit de telle sorte que sa capacité, commandée par Q_1 , varie de 60 pF à 0 V à 5 pF à la tension de pincement, ce qui permet à Q_3 de travailler dans les meilleures conditions.

5. LES BIFET

Alors que la technologie BIMOS réunit, sur un même substrat, des transistors bipolaires et des MOS, la BIFET associe des bipolaires et des JFET. On a utilisé de telles désignations dès lors que l'on a réalisé des amplificateurs opérationnels très performants, bien que la BIFET pré-existait... C'est probablement *RCA* qui a donné le coup d'envoi des BIMOS et *National Semiconductor* celui des BIFET.

Ces deux technologies seront illustrées au chapitre des applications. On constatera qu'elles mènent à des produits de très haute qualité, capables de battre les classiques 741 (bipolaires); peut-être seront-elles à même de dominer le marché. En effet, les entrées de tels amplificateurs opérationnels se font sur des transistors unipolaires, donc à haute impédance et très faible consommation; pour les BIMOS, les sorties se font sur MOS complémentaires; dans les deux cas, l'amplification est assurée en bipolaire. Les composants aux performances optimales sont ainsi employées en chaque point de la chaîne.

La figure 23 illustre le processus de fabrication d'une structure BIFET, avec un JFET et un NPN, selon *Texas*.

CHAPITRE VII

LES CIRCUITS A TRANSFERT DE CHARGES : CCD, BBD, ET VARIANTES

Les circuits à transfert de charges, ou CTD (de charge transfer devices) constituent une catégorie nouvelle d'éléments, propre aux circuits intégrés, car sans précédent ni équivalent en discret. Leur structure les apparente aux MOS et à leur mode de fonctionnement ; leurs performances sont assez extraordinaires. Les CTD comprennent essentiellement :

- 1. Les CCD (charge coupled devices), les plus simples.*
- 2. Les BBD (bucket brigade devices), ou éléments à chapelets.*
- 3. Les SCT (surface charged transistor), en variante des précédents.*
- 4. Les CID, ou circuits à injection.*

Ils s'appliquent aux registres et aux mémoires à très haute densité d'intégration, au numérique et à l'analogique (lignes à retard, filtres, multiplexeurs), et aux dispositifs de prise de vue à l'état solide, jusqu'à la caméra TV.

1. PRINCIPE DES CCD

La structure d'un CCD est extrêmement simple (fig. 1) : sur un substrat de silicium uniformément recouvert d'oxyde, on a déposé des électrodes métalliques semblables à celles des gates de MOS. Ainsi :

- il n'y eu aucune diffusion;
- on n'a ouvert aucune fenêtre dans l'oxyde;
- les seuls masquages visent les métallisations et la distribution de l'oxyde, en couches minces et épaisses;
- aucun souci d'isolement n'est à respecter;
- par conséquent, le CCD se révèle économique et permet une grande densité d'intégration.

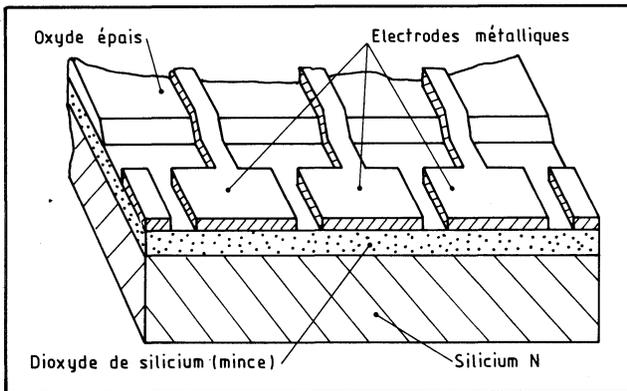


Fig. 1. — Structure d'un CCD.

Comment s'utilise-t-il? Supposons que l'on injecte dans le substrat des porteurs minoritaires positifs, et cela d'une façon quelconque non indiquée sur les schémas (par exemple, via un transistor classique intégré sur le même substrat, ou par bombardement photonique...). Si l'une des électrodes métalliques est polarisée négativement, au-delà de la tension de seuil, il est évident qu'elle va attirer ces charges positives dans le substrat sous-jacent. Ces charges y resteront piégées (fig. 2 a).

Si, ensuite, on polarise encore plus négativement l'électrode de commande voisine, on conçoit que ces mêmes charges soient alors attirées par celle-ci et se déplacent dans la nouvelle zone du substrat (b).

Ainsi, une suite de cellules CCD agit comme un registre à décalage. Dans l'exemple ci-dessus, trois tensions, dont la plus faible ne sert qu'à la création d'un canal minimal, ont été nécessaires; on dit qu'il s'agit d'une *commande triphase*, illustrée par la figure 3. En appliquant successivement ces tensions à un groupement de trois éléments, on provoque le transfert des charges d'un élément dans l'autre avec les étapes suivantes :

- en a, les éléments 1, 4 et 7 sont activés et piègent les charges;
- en b, on déplace les charges vers 2, 5 et 8 par une tension V_3 plus négative que V_2 ;
- en c, on ramène la polarisation de 1, 4 et 7 à la plus faible tension, et celle de 2, 5 et 8 à la tension négative moyenne. Ainsi, il y a eu un cycle complet de transfert en trois étapes entre un élément et son voisin. On pourrait continuer de la sorte...

Ce fonctionnement est donc du type *dynamique*; la durée du stockage des porteurs minoritaires, limitée par la création des paires électron-trou par effet

Fig. 2. — Le CCD en fonction mémoire (a) et registre à décalage (b) : seule la polarisation appliquée aux électrodes diffère.

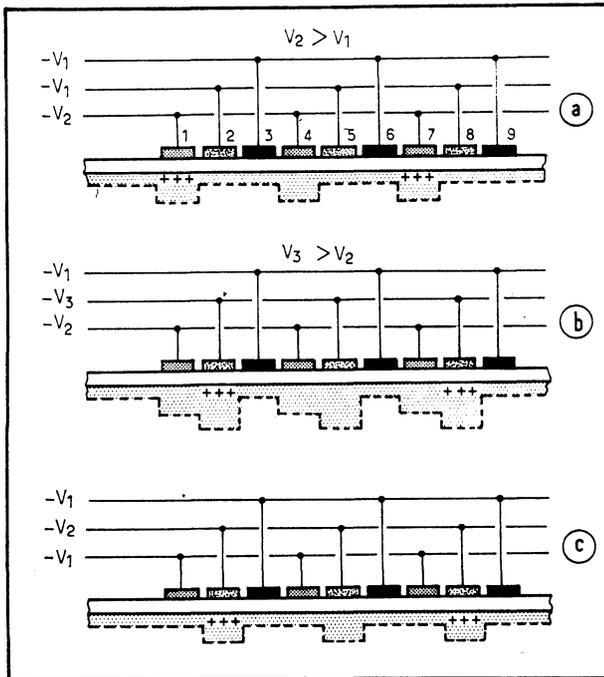
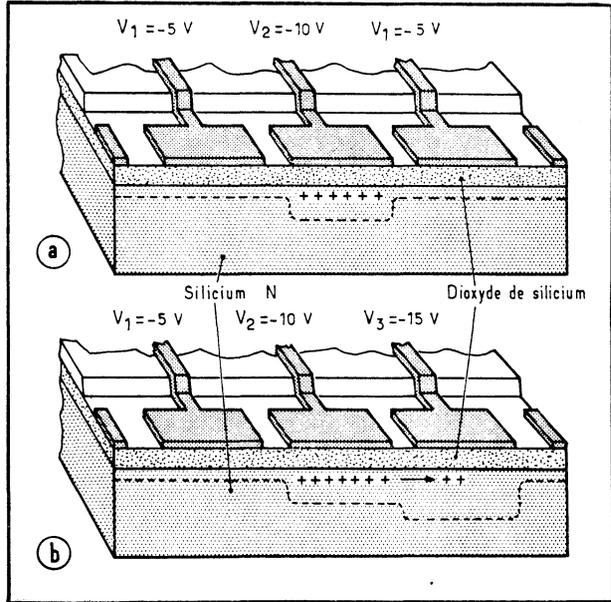


Fig. 3. — La commande triphase des CCD.

thermique, pourrait atteindre la seconde. La vitesse de transfert des charges permet de travailler jusqu'à 20 MHz pour une consommation de $5 \mu\text{W/bit}$. Un bit, soit trois éléments consécutifs, occupe $1\,250 \mu\text{m}^2$ de silicium.

De tels CCD ont été annoncés dès 1970 par W. S. BOYLE et G. E. SMITH, des *Bell Laboratories*. Si l'on compare leur structure à celle du transistor MOS ou bipolaire, on sera convaincu de leur simplicité (fig. 4).

2. LES BBD

Les BBD ne diffèrent des CCD que par la diffusion de régions P dans le silicium, ce qui les apparente fortement aux MOS (fig. 5 a). Ce sigle BBD provient de « bucket brigade devices », qu'on peut traduire par éléments à chapelets ou « chaînes à seaux », car il évoque les anciennes méthodes de lutte contre l'incendie,

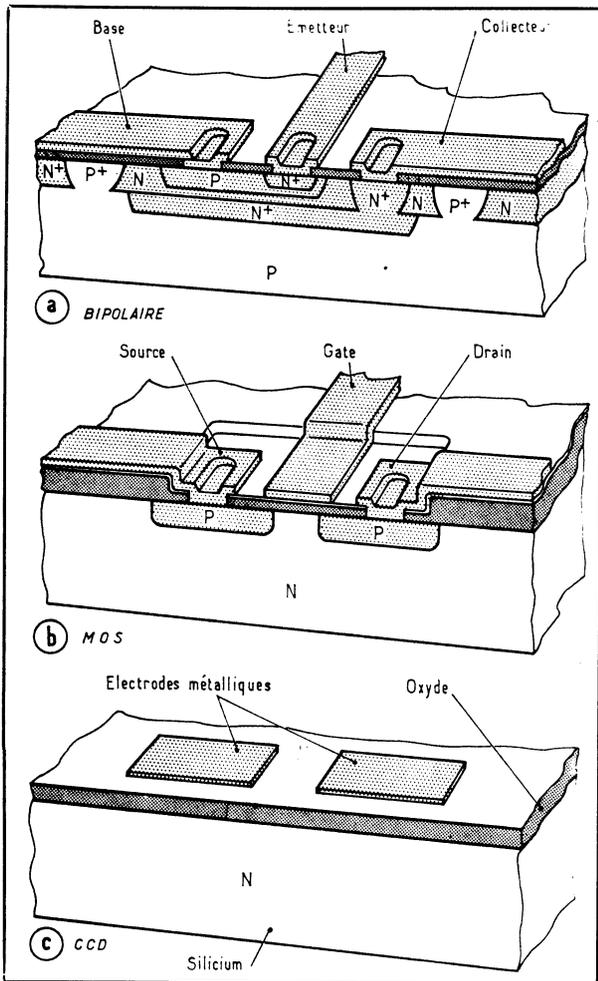
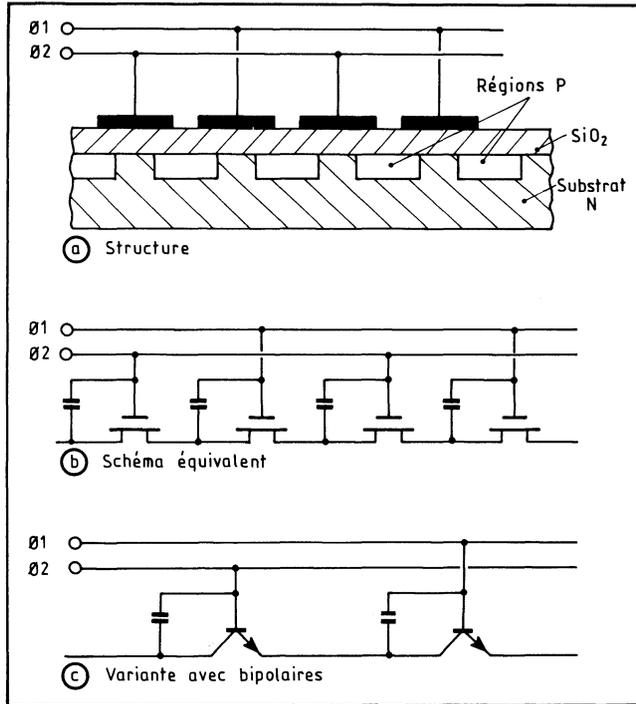


Fig. 4. — Structures comparées de transistors bipolaire (a), MOS (b), et d'un élément CCD tel que conçu par les *Bell Telephone Laboratories* (c). Dans ce dernier cas, le processus de fabrication est évidemment simplifié à l'extrême.

Fig. 5. — Principe des éléments « à chapelet », ou BBD (« bucket brigade devices »); ils se différencient des CCD par une zone diffusée P qui les rapproche des MOS (a). Schéma équivalent aux BBD (b). Structure BBD à transistors bipolaires (c).

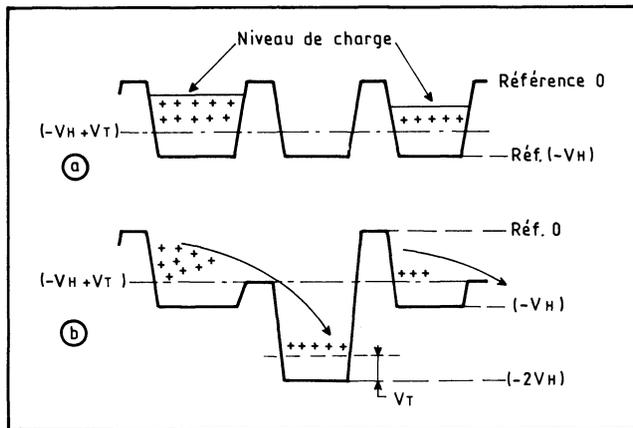


où l'on se passait des seaux d'une main à l'autre... Les BBD sont antérieurs aux CCD puisqu'ils étaient annoncés, en 1969, par F. L. S. SANGSTER et K. TEER, des laboratoires Philips de Eindhoven (Hollande).

Le schéma équivalent à cette structure est donné dans la même figure (b). En variante (c), on voit qu'on pourrait même utiliser des bipolaires au lieu de MOS.

Le fonctionnement, identique dans son principe à celui des CCD, se suffit cependant ici de deux phases. On peut l'illustrer comme le montre la figure 6.

Fig. 6. — Les charges (analogiques) positives, accumulées en a, s'écroulent dans le « puits » créé en (b) en doublant le potentiel négatif de référence.



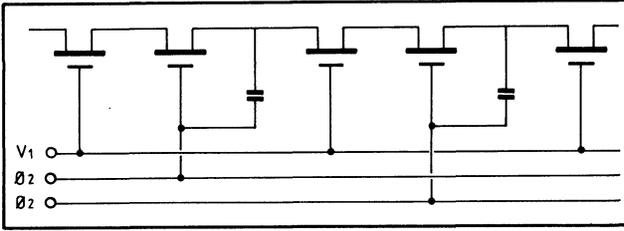


Fig. 7. — Variante des BBD : deux MOS assurent une meilleure séparation des « puits », et favorisent l'aptitude des circuits à travailler en analogique (schéma Philips).

Si V_H représente le niveau du signal d'horloge et V_T la tension de seuil dans le cas des MOS (le niveau V_H étant négatif par rapport à la masse), le niveau de référence sera $-V_H + V_T$, et la charge emmagasinée lui sera superposée. Dans la courbe en *a*, le second élément ne contient aucune charge; par contre, les premier et troisième sont chargés à des niveaux différents (ce qui confirme leur aptitude à traiter des signaux analogiques). En *b*, la tension d'horloge a été doublée et est passée à $-2 V_H$: les charges stockées par l'élément précédent vont « s'écrouler » dans le « puits » de l'élément polarisé à $-2 V_H$. On retrouve ainsi les deux modes de fonctionnement : mémoire et registre.

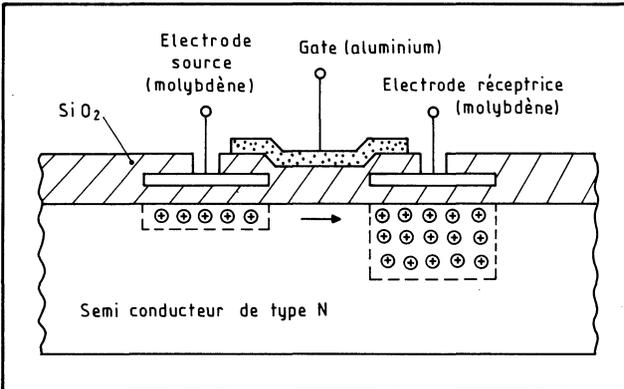


Fig. 8. — Structure d'un SCT.

La charge et la décharge constituant des phénomènes non linéaires, les niveaux de charge se détériorent rapidement; aussi, les éléments à transfert de charge sembleraient-ils plus adaptés au numérique. Pourtant, on peut, en séparant mieux les capacités, créer des circuits traitant de façon très satisfaisante les signaux analogiques. Cela est obtenu avec la structure à double transistor de la figure 7.

Les BBD (ou BB, en abrégé...) consomment également $5 \mu\text{W/bit}$ mais leur vitesse est moindre : 10 MHz. Par contre, ils traitent mieux les signaux analogiques. Leur densité d'intégration est de l'ordre de $1\,600 \text{ bits/mm}^2$.

Ainsi, réalisée par Philips, une ligne à retard expérimentale en BB à transistors bipolaires, comportant 72 étages et traitant des signaux vidéo, occupe une puce de

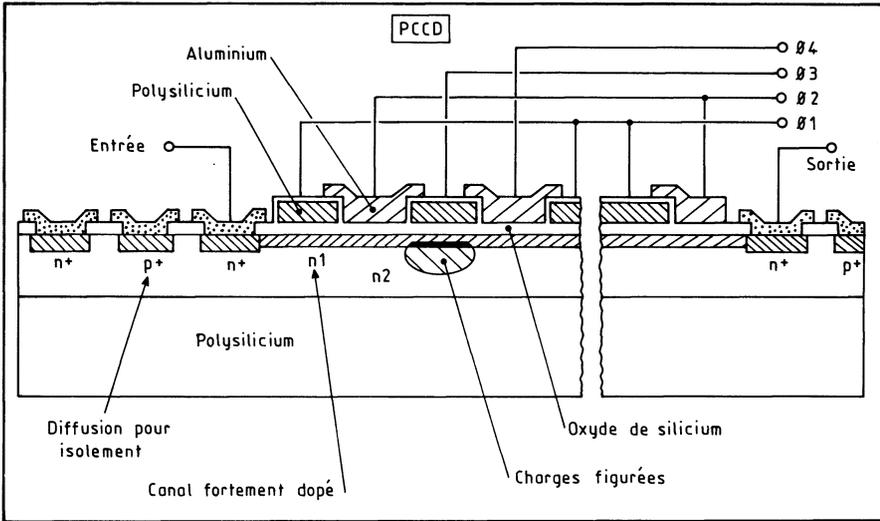


Fig. 9. — Coupe d'un PCCD expérimental. Les charges circulent de la gauche vers la droite.

silicium de $150 \mu\text{m}^2$; un réseau BB de 512 étages MOS occupe 9 mm^2 , pour une capacité-mémoire de 3 pF. Les CCD occuperaient une surface trois fois moindre, et permettraient le regroupement de 10 000 bits sur une puce.

3. LES SCT

Une variante des CTD est constituée par les SCT, ou *transistors à charge de surface*, mis au point par *General Electric* (fig. 8). Comme dans les CCD, des régions de déplétion sont formées sous les électrodes. Dans ces régions sont stockés des porteurs minoritaires; le transfert de ces charges est commandé par l'électrode centrale en aluminium. *General Electric* a dénommé « réservoirs » les régions de déplétion.

4. LES PCCD

Les deux limitations dont souffrent les CCD sont :

- une vitesse de transfert limitée;
- surtout, une perte de charge à chaque transfert.

Interdépendantes, ces deux caractéristiques de vitesse et de perte de charge résultent d'un compromis : si l'on réduit la fréquence d'horloge, il est certain qu'on aura le temps de transférer un taux plus élevé de la charge...

Affrontant ce problème, L. J. M. ESSER, des Laboratoires de recherches de *Philips*, à Eindhoven (Hollande) a créé, en 1973, un circuit à transfert de charge amélioré qui résulte de l'observation suivante : la vitesse à laquelle la charge stockée par un condensateur est transférée à un élément voisin dépend principalement de la vitesse avec laquelle la dernière fraction utile de la charge est transférée.

Ce nouveau circuit, nommé *peristaltic charge coupled device*, soit *PCCD*, est représenté figure 9. Il est conçu de façon que la dernière fraction de charge se trouve dans la masse même du silicium N, car la mobilité des porteurs est plus grande au sein du matériau qu'au voisinage de la surface; cette dernière fraction se trouve en outre, accélérée par un fort champ électrique. Quant au reste de la charge, c'est-à-dire l'essentiel, elle se trouve proche de la surface et son transfert n'exige qu'une faible tension de commande.

La manière dont la charge d'espace, ce « nuage » de charge, est propulsé d'une cellule à l'autre fait songer à un mouvement péristaltique, d'où le nom donné à ce circuit (1).

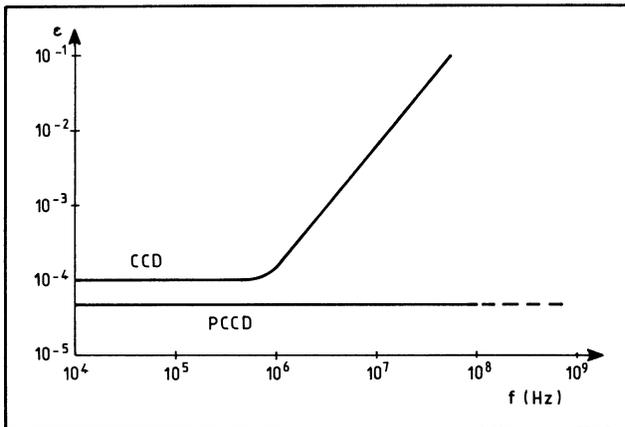


Fig. 10. — « Inefficacité » comparée d'un PCCD et d'un CCD « classique ». Elle est notée, ici, pour un transfert d'une cellule à l'autre en fonction de la fréquence d'horloge.

Comme on le voit sur la figure, un canal ininterrompu N a été réalisé par diffusion ou implantation ionique dans le silicium P. Les charges sont introduites, à gauche, via la jonction métal-silicium N^+ . Les électrodes sont soumises à des signaux d'horloge déphasés, par exemple, de 90° , qui propulsent ces charges vers la droite où elles sont récupérées par l'électrode de sortie.

Un PCCD expérimental, conçu par M. J. J. THEUNISSEN, également du laboratoire de recherches *Philips*, s'est ainsi révélé capable d'un rendement de 99,99 % à une fréquence d'horloge supérieure à 100 MHz; la perte de charge au cours d'un transfert n'est donc plus que de 0,01 %, ce qui est important puisque l'on pourra, ainsi, recourir à des transferts sur plus de 1 000 pas successifs.

La courbe de la figure 10 compare le rendement d'un PCCD et d'un CCD, à partir de la notion inverse de l'efficacité, c'est-à-dire de l'inefficacité notée ϵ . La figure montre le progrès important réalisée par le PCCD qui pourrait, à l'avenir, travailler jusqu'à 1 GHz.

1. Péristaltique adj. (du préf. *péri*, et du grec *stallein*, serrer). Se dit du mouvement par lequel l'œsophage et les intestins se contractent sur eux-mêmes et favorisent la déglutition, puis la digestion. (*Le Petit Larousse*.)

5. PRINCIPES DES APPLICATIONS

Les dispositifs à transfert de charge ont trouvé trois grands domaines d'application :

1. *Les registres à décalage et les mémoires dynamiques.*
2. *En linéaire et numérique, les lignes à retard, le multiplexage et le filtrage.*
3. *L'optoélectronique jusqu'à la caméra de prise de vues à l'état solide.*

On va en tracer ici les grandes lignes en raison de la place particulière qu'occupent ces éléments, qui ne seront ensuite cités que pour mémoire essentiellement.

6. REGISTRES ET MÉMOIRES

L'intérêt des CCD appliqués aux registres et mémoires, surtout, est double :

1. La densité d'intégration est considérablement plus élevée.
2. Le prix de revient au bit est plus faible.

Parmi les premières mémoires commercialisées figurent la 450 de *Fairchild*, de 9 K bits, dont la structure est donnée figure 11, et la 2416 de *Intel*, de 16 K, toutes deux à organisation série. Puis, en 1977, trois fabricants ont proposé des

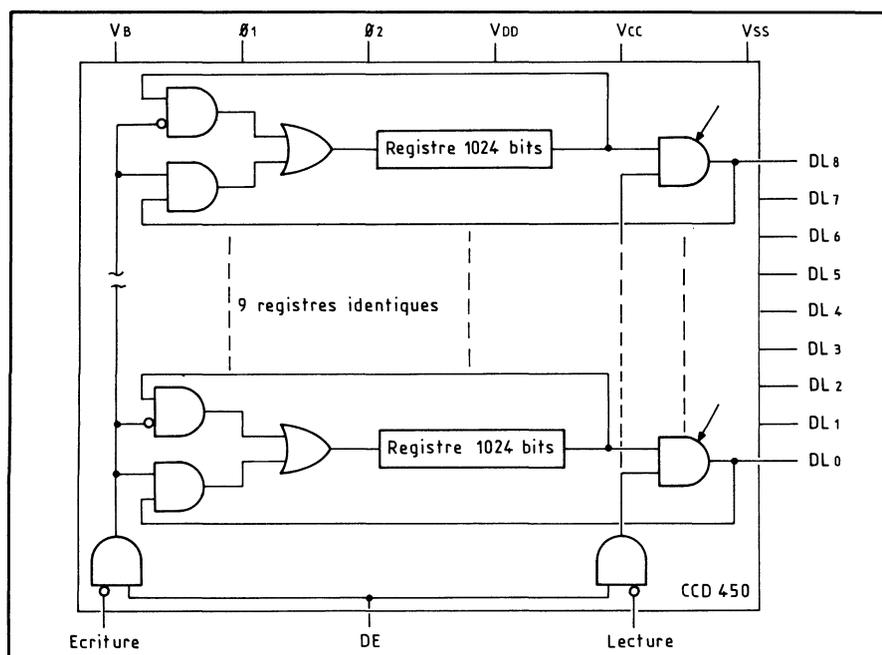


Fig. 11. — Le CCD 450 Fairchild de 1 024 mots de 9 bits.

64 K : *Intel* avec la 2464, *Texas* avec la TMS 3064 et *Fairchild* avec la 464, déjà bien plus intéressantes à divers titres.

Ainsi, la 464 tout comme la 3064 est organisée en 16 registres adressables de 4 K bits. La fréquence maximale de transfert est de 5 MHz et de ce fait, l'accès moyen se fait en quelque 400 μ s. La structure interne est du type « série-parallèle-série » (ou SPS) : un registre d'entrée série alimente les registres 4 K en parallèle, lesquels à leur tour attaquent un registre série de sortie. Pour sa 2464, *Intel* a choisi une organisation en 256 registres adressables de 256 bits.

La consommation typique de la 3064 de *Texas* est de 300 mW à 5 MHz, en activité. Cette consommation peut être divisée par dix au repos et ramenée à 30 mW.

Il est certain qu'on n'en restera pas là et que, les densités d'intégration croissant, les mémoires CCD concurrenceront les RAM classiques, en particulier en mémoires intermédiaires en informatique, et dans le remplacement éventuel des disquettes magnétiques pour lequel les idées se sont multipliées ces dernières années. Voyons-en le principe.

Le disque en CCD

L'une des toutes premières sociétés à proposer la réalisation de l'équivalent du disque magnétique a été *RCA*, avec des registres CCD de 8 K bits. Mais il s'agit d'un substitut sans aucune pièce mécanique en mouvement.

Cette formule, qui avait été envisagée précédemment avec des registres MOS, offre un net avantage économique avec les CCD qui simulent les pistes du disque (fig. 12). Plusieurs CCD étant montés en parallèle, on aboutit alors à l'équivalent d'un tambour, ou d'un minidisque magnétique. L'électronique de commande a été réalisée en COSMOS; l'horloge est biphase. On reviendra sur ces thèmes dans les chapitres dédiés aux mémoires.

Pour justifier cette voie, *RCA* avait recherché une forte densité d'intégration; dès 1974, la firme logeait une cellule complète sur moins de 0,001 mm², les composants étant séparés de 2 à 3 μ m, avec des canaux de 1 μ m. Des surfaces de moins de 500 μ m² ont été obtenues depuis. La perspective est de réaliser des circuits de 500 000 à 1 million de bits sur une puce.

Ainsi, et selon B. AUGUSTA, d'*IBM*, il serait possible de réaliser une mémoire de masse de 100 mégabits avec des circuits CCD. Le temps moyen d'accès à l'information serait inférieur à 0,5 milliseconde, et la vitesse de transfert des données serait supérieure à 32 mégabits par seconde.

7. LES LIGNES A RETARD

Les registres à décalage se transposent parfaitement en lignes à retard; la fréquence d'horloge qui les commande pouvant varier à volonté, le retard est ajustable. De plus, on peut aussi bien envisager le traitement d'informations numériques qu'analogiques.

Ainsi, F.L.J. SANGSTER, de *Philips*, a été l'un des premiers à proposer l'emploi de lignes à retard BBD pour la correction d'erreurs de temps dans la reproduction de signaux audio et vidéo enregistrés sur bande magnétique. Ces erreurs sont dues

à l'étirement de la bande et à des variations de la vitesse de défilement; elles sont corrigées en injectant le signal à corriger dans une ligne à retard BBD, en utilisant les erreurs pour la commande de la fréquence de décalage. On peut encore utiliser une variation de la fréquence de décalage pour générer délibérément des distorsions de l'axe du temps du signal, l'étaler, ou le comprimer.

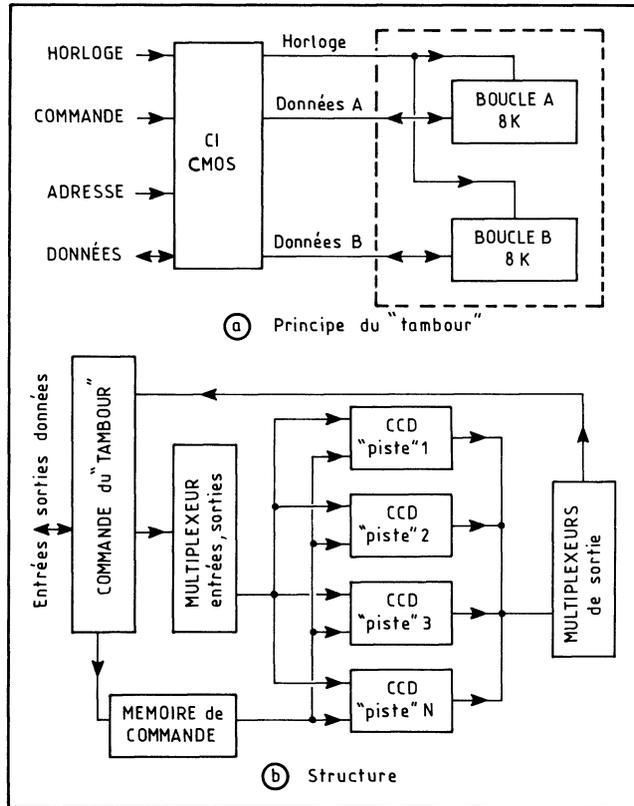


Fig. 12. — Principe de la « cyclomémoire » équivalente à un disque magnétique.

Une application pratique de ce dernier principe se situe en téléphonie, où il est souhaitable de transmettre un certain nombre de signaux à bande étroite dans un canal à large bande. Dans ce but, chaque signal est stocké dans une ligne à retard BBD, correspondant à leur petite largeur de bande, (par exemple : 10 kHz), puis lu à grande vitesse (par exemple : 1 MHz). Un fragment de signal de 5 millisecondes peut, ainsi, être comprimé en 50 microsecondes. Plusieurs canaux sont lus en séquence, donnant naissance à un signal multiplexé comprimé que l'on transmet vers le récepteur où le processus inverse est effectué.

Une autre application des circuits BBD, dans le domaine audio, est le codage de conversations téléphoniques. La voix est échantillonnée, et les échantillons sont injectés, de façon programmée, dans une batterie de lignes à retard BBD.

Le circuit « chaîne à seaux » peut aussi être utilisé pour la réalisation de filtres transversaux : le signal injecté traverse un certain nombre d'éléments de retard et

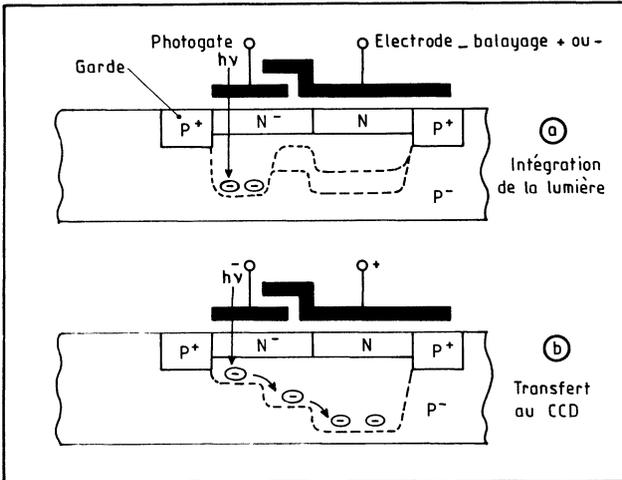
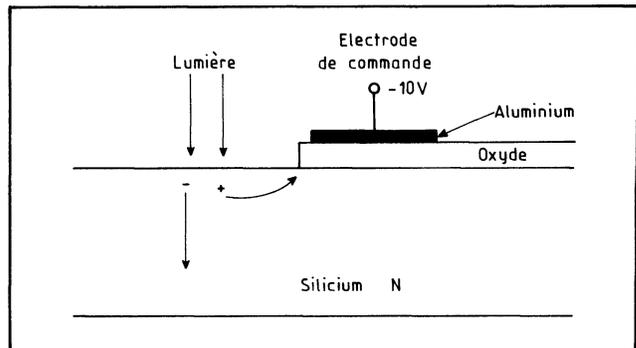


Fig. 13. — Principe des applications opto-électroniques.

se trouve pondéré par un facteur qui dépend du retard enregistré. Si un élément de circuit apporte un retard T_r sur la tension $V_e(t)$ appliquée au temps t à l'entrée du filtre, le retard après passage du signal dans k éléments sera : kT_r . Ainsi, le signal issu du $k^{\text{ième}}$ élément, soit $V_k(t)$, correspondra au signal injecté à l'entrée du filtre au temps $t - kT_r$, soit $V_e(t - kT_r)$. En chaque élément k de retard, le signal V_k est pondéré par un coefficient h_k , qui est, bien entendu, fonction du retard appliqué : $h_k = h(kT_r)$. Enfin, le signal de sortie du filtre sera la sommation de tous les signaux V_k , pondérés par les facteurs h_k ; ce sera donc une somme de termes du type $h(kT_r) \cdot V_e(t - kT_r)$, k prenant toutes les valeurs entières inférieures ou égales au nombre d'éléments de retard du filtre. On opère ainsi une convolution sur le signal d'entrée. De tels filtres ont été étudiés par *Texas Instruments*; ils servent de filtres adaptés pour la détection de signaux noyés dans des bruits de fond importants (systèmes militaires de communication, radars,...). Il a été démontré que ces filtres adaptés peuvent également être réalisés en technologie CCD.

Fig. 14. — La cellule de base des photo-CCD.



Dès 1974, le « père » du système PAL de télévision couleur, Walter BRUCH, de *AEG-Telefunken*, démontrait de nombreux systèmes de correction d'images et de sons, de compression et d'expansion des informations, etc., réalisés avec des CCD sur les principes énoncés ici.

Philips et *Matsushita* développaient, fin 1974, le MN 3001, de 2 fois 512 étages, sur une puce de $2,8 \times 3,68$ mm, travaillant entre 10 kHz et 800 kHz et destiné à des applications audio : élimination des fluctuations de vitesse de magnétophones, effets de trémolo ou autres, trucages, compression-expansion en téléphonie... La commande est biphasée et la structure est celle des BBD avec MOS à double gate.

À la fin de 1975, *Philips* annonçait le TDA 1022, ligne à retard BBD de 512 bits recourant à la fois à une technologie à MOS complémentaires et à la « LOCMOS », capable de retarder des signaux de 51,2 ms à 512 μ s. Il accepte 2,5 V_{eff} à l'entrée et son atténuation typique n'est que de 3,5 dB. La fréquence d'horloge peut varier de 5 à 500 kHz. Cet ensemble de caractéristiques en faisait certainement le BBD le plus performant de cette époque.

8. LES APPLICATIONS OPTO-ÉLECTRONIQUES

Dans les applications opto-électroniques, on exploite la photo-sensibilité du silicium pour réaliser des dispositifs :

1. *Linéaires*, ou
2. *Plans*.

Dans les deux cas, le principe en est le suivant (fig. 13). Les photons viennent frapper le silicium et créent des paires électron-trou; si l'électrode de commande est polarisée négativement (pour un substrat N), les charges minoritaires P créées vont être récupérées dans la zone d'inversion située sous l'électrode de commande, et l'on disposera ainsi de charges qui témoigneront de la quantité de lumière reçue.

L'intérêt ainsi manifesté envers les circuits CCD s'explique par le fait que ceux-ci devraient rapidement être, sous de nombreux aspects techniques et économiques, aussi performants que les tubes vidéo.

Les structures réellement réalisées sont toujours un peu plus complexes. Par exemple, *Fairchild* utilise la cellule dessinée figure 14. Les photons traversent la mince électrode notée *photogate* et viennent frapper le silicium, créant des paires électron-trou (en *a*); l'électrode de balayage se trouve alors indifféremment au *plus* ou au *moins*; en une seconde étape, le photogate est rendu négatif alors que l'électrode de balayage devient positive (en *b*) : les charges négatives, puisque le substrat est ici P, sont transférées sous l'électrode de balayage, entrée du circuit à transfert de charge. Le traitement ultérieur ne dépend plus que des principes retenus.

Les méthodes d'analyse

La zone photosensible du silicium, constituée par une jonction NP, peut être assimilée à une photodiode. Si l'on suppose que celle-ci est suivie par un MOS de commande, une chaîne constituée par de telles cellules pourra être représentée comme le montre la figure 15 : la charge accumulée par les photodiodes est transférée, à tour de rôle, vers la ligne de sortie par des MOS-interrupteurs, commandés séquentiellement grâce à un ordre circulant le long d'un registre CCD. Dans le schéma, et si le balayage est assuré à une fréquence télévision, on obtient directement une sortie vidéo.

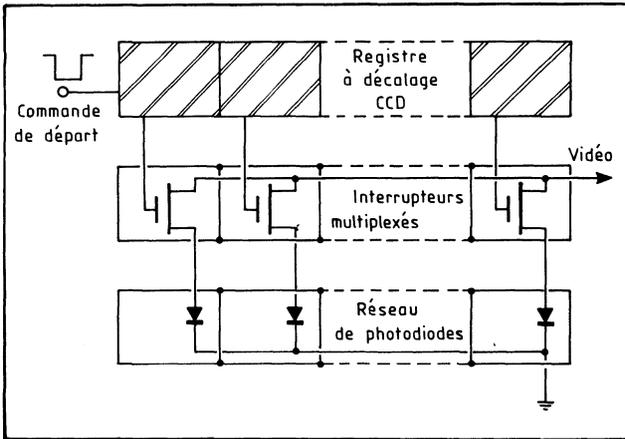


Fig. 15. — Dispositif classique d'analyse d'une ligne d'image. Les interrupteurs MOS sont commandés séquentiellement.

Une autre méthode, proposée par *Philips*, consiste à transférer simultanément toutes les charges au registre à décalage BBD qui fournit la vidéo à sa sortie (fig. 16).

L'amplificateur à gate flottant

Le courant fourni par la photodiode est très faible. Pour l'exploiter au mieux, de nombreux principes ont été élaborés; ainsi, pour une amplification directe du signal de sortie, il faut disposer d'un excellent amplificateur. L'une des solutions les plus originales est celle de *Fairchild*, qui a réalisé un amplificateur à gate flottant, appelé FGA (*floating gate amplifier*).

Fig. 16. — Principe adopté par Philips pour analyser l'image. Ici, tous les interrupteurs sont commandés simultanément; à ce moment, le défaut de charge des diodes est transféré au circuit à chapelet.

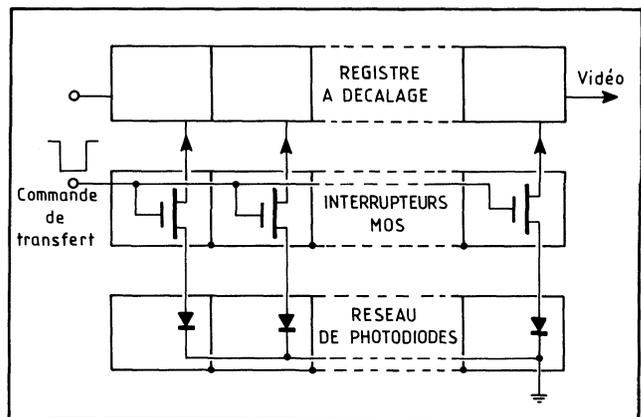
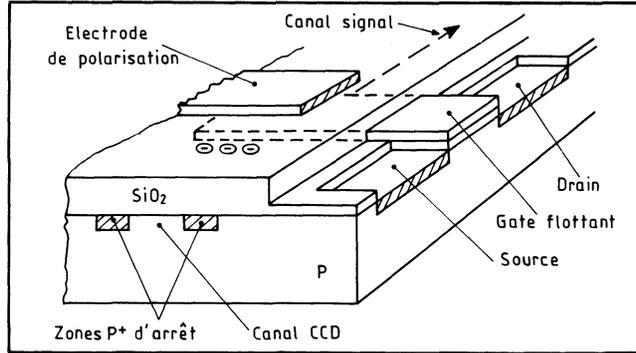


Fig. 17. — Principe de l'amplificateur à gate flottant, ou FGA.



L'idée retenue est la suivante (fig. 17) : le photocourant dans le silicium est couplé capacitivement à une électrode totalement noyée dans l'oxyde isolant; or cette même électrode, dans son prolongement, sert de gate à un MOS classique!

Puisque le signal original utile n'est pas détruit par ce processus, il peut même resservir une seconde fois, ou de multiples fois, ce qui constitue une astuce supplémentaire. On en arrive alors à un montage tel que celui de la figure 18 qui garantit effectivement, selon *Fairchild*, un excellent rapport signal sur bruit pour une même bande passante.

Il va de soi que l'électrode flottante doit être très proche du canal du MOS, alors que l'électrode normale du CCD doit en être éloignée via de l'oxyde épais.

L'analyse à décharge de photodiode

Une autre méthode consiste non plus à mesurer un courant de charge, mais un courant de décharge de la photodiode élémentaire. Pour cela, la photodiode est

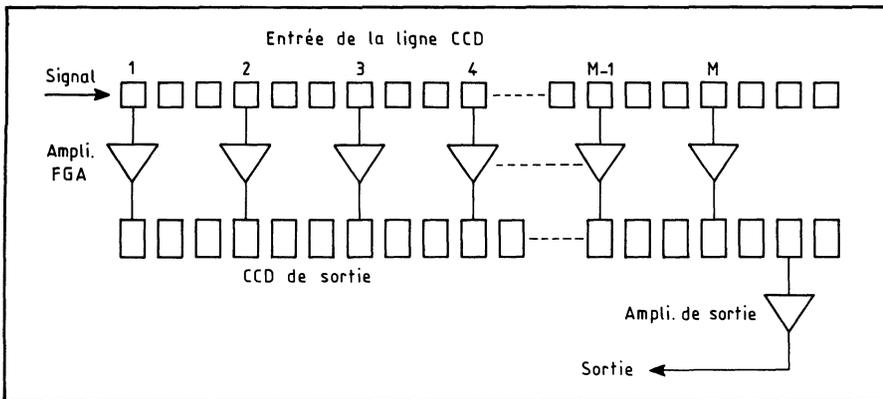


Fig. 18. — Une idée astucieuse d'utilisation successive des mêmes charges avec des FGA.

chargée au préalable à un niveau de référence, c'est-à-dire qu'on lui applique tout simplement une polarisation inverse; sous l'effet de la lumière incidente, elle va se décharger entre deux interrogations successives; chaque interrogation consistera alors à lui réappliquer le niveau inverse de polarisation de référence et à lui fournir un courant de recharge qui témoignera de la quantité de lumière reçue.

L'analyse linéaire

Le réseau d'analyse linéaire est constitué par une série de cellules élémentaires. Ainsi, dans le circuit CCD 110 de *Fairchild*, en technologie Isoplanar, on trouve 256 cellules réalisées comme le montrait l'une des figures précédentes.

Pour une plus grande densité, les circuits CCD se répartissent de part et d'autre du réseau photosensible, comme le montre la figure 19. Une commande à deux phases suffit alors, ce qui laisse 50 % du temps d'intégration de la lumière pour chaque cellule.

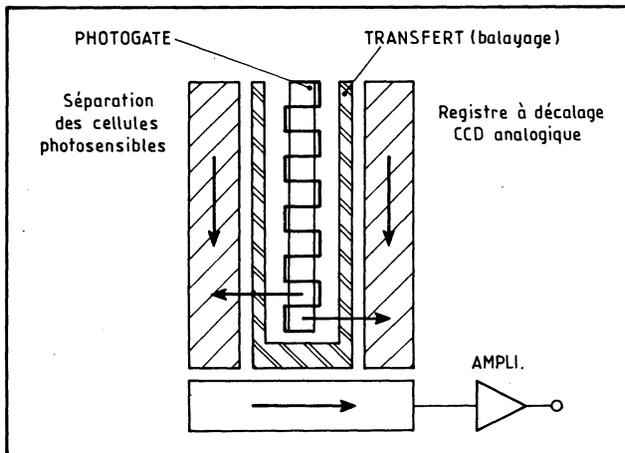


Fig. 19. — Principe de l'analyse linéaire.

Le synoptique réel du CCD 110 est donné figure 20. Les photocellules sont numérotées de 1 à 256 (au centre); leur électrode de commande est P (*photogate*). Le boîtier comporte, bien entendu, une fenêtre transparente à la lumière.

Le fonctionnement est le suivant : les photons traversent les photogates, en silicium polycristallin transparent, et sont absorbés par le silicium en créant des électrons et des trous. Les électrons restent piégés et s'accroissent linéairement avec l'éclairage et le temps. La charge est donc analogique.

Les charges de chaque cellule sont transférées alternativement vers les registres de droite et de gauche sur commande ϕ_{XA} et ϕ_{XB} , puis déplacées par paquets, en série, vers le détecteur-préamplificateur via 130 cellules CCD (de part et d'autre). Elles sont alors appliquées au gate du MOS canal N de sortie, polarisé via une diode en fonction du niveau du signal.

Pour une fréquence de décalage maximale de 5 MHz, la fréquence d'analyse atteint 10 MHz; les tensions de commande restent inférieures à 20 V.

Les applications d'un tel circuit concernent, en particulier, la lecture optique et tous les systèmes d'analyse à haute vitesse.

L'analyse d'image

L'analyse d'image qui mène, elle, à la caméra TV totalement à l'état solide, s'est traduite par de nombreuses réalisations. En 1975, les premiers systèmes compatibles avec la définition télévision sont ainsi apparus, à titre expérimental.

Par exemple, RCA a réalisé un circuit de 163 840 cellules CCD distribuées en réseau de 512×320 éléments et fonctionnant en 525 lignes avec une horloge

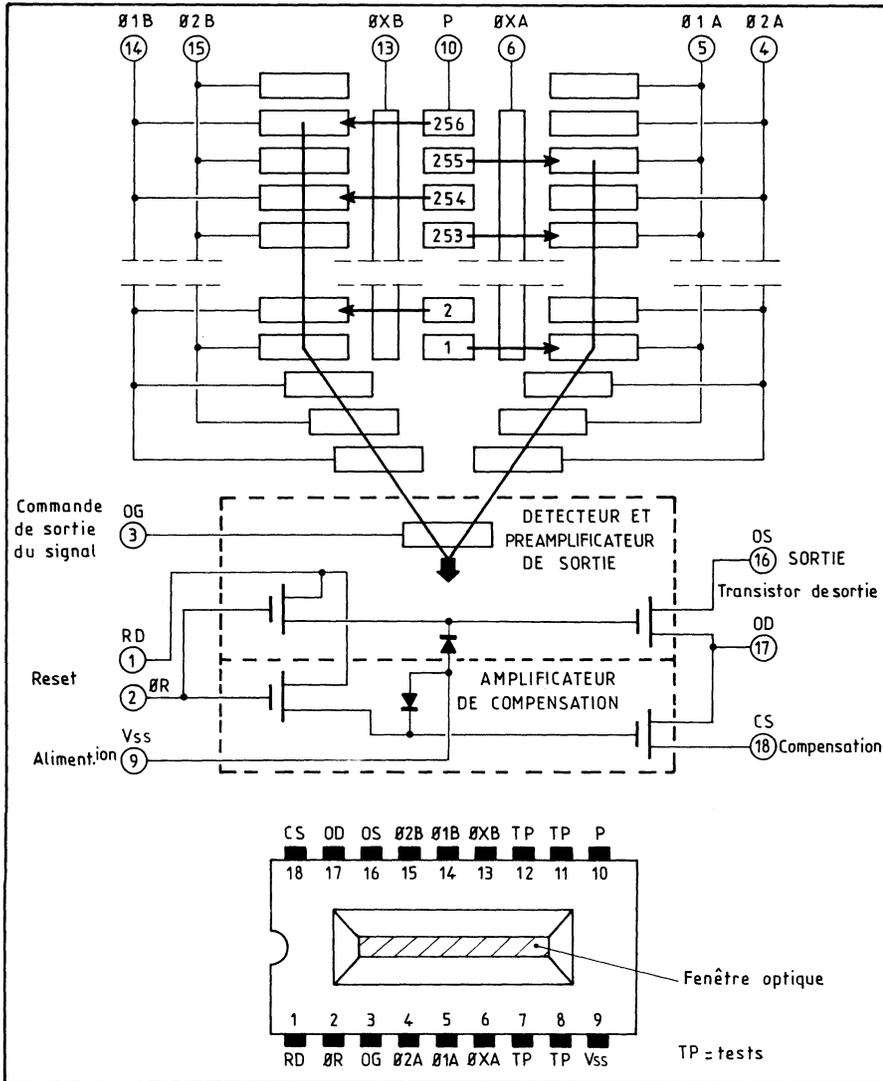


Fig. 20. — Synoptique du CCD 110, de 256 points sensibles (a); présentation (b).



Fig. 21. — Image obtenue avec un réseau expérimental de 100 par 100 points d'analyse.

d'environ 6 MHz; la bande passante est de 3 MHz. Ce réseau CCD, ou plus exactement SID (de *silicon imaging device*), mesure moins de 15×20 mm, et il a été monté dans une caméra minuscule, la TC 1150 dont l'objectif à focale variable de 14 à 38 mm constitue la partie la plus volumineuse.

Prenons l'exemple d'un réseau plus simple, celui du CCD 201 de *Fairchild*, constitué par un ensemble de 10 000 cellules réparties en réseau de 100×100 . La surface occupée par une cellule est d'environ 30×20 microns; leurs centres sont espacés verticalement de 30 microns, horizontalement de 40 microns.

En raison des deux phases d'horloge, l'analyse est entrelacée sans problème, et la fréquence de décalage peut être assurée à la fréquence normale de balayage lignes, ce qui simplifie d'autant les circuits périphériques. La fréquence maximale d'horloge pour l'analyse horizontale est de 4 MHz; la consommation globale est de l'ordre de 50 mW.

Il est évident que ce réseau est relativement peu important mais tel quel, il se prête déjà à la réalisation de caméras de surveillance, par exemple. La figure 21 montre le type d'image qu'il permet d'obtenir.

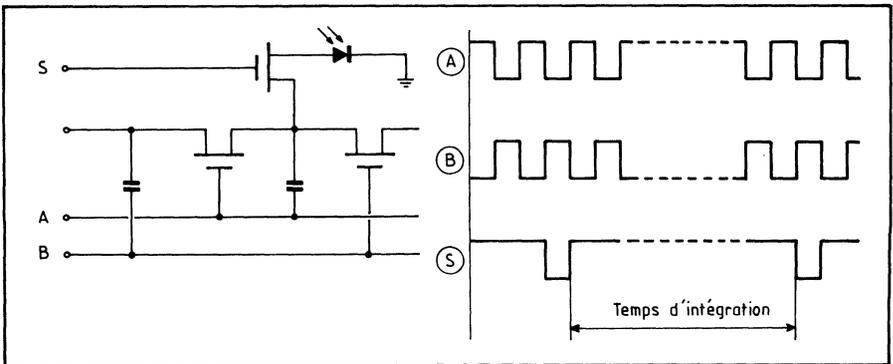


Fig. 22. — Schéma détaillé d'un élément BBD. La photodiode, polarisée en inverse, perd sa charge entre deux interrogations. Cette perte de charge est transférée, via le MOS interrupteur, au réseau BBD proprement dit.

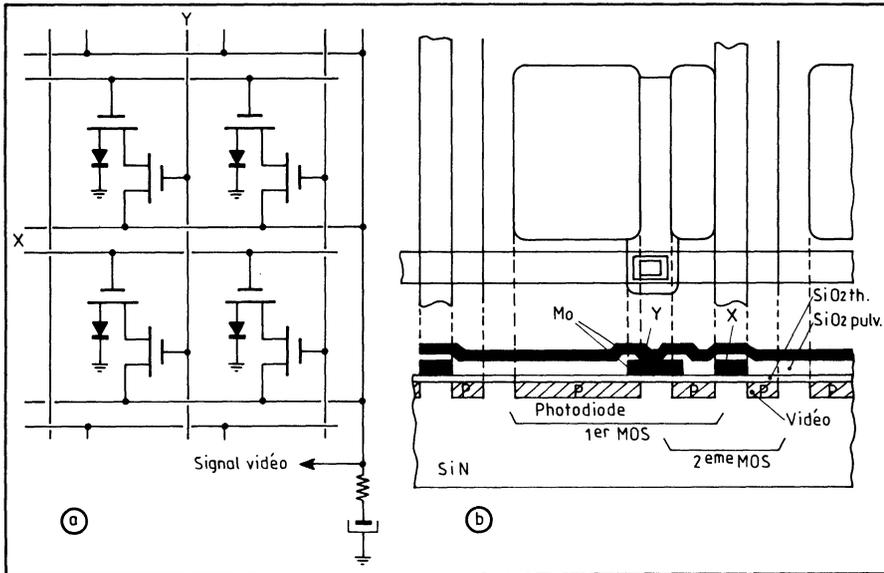


Fig. 23. — Principe et structure du point image de la matrice de prise de vues du LEP, à RMOS.

Les laboratoires *Philips* de Eindhoven ont retenu un schéma différent (fig. 22), à BBD. Le réseau de photodiodes intègre la lumière incidente et le déficit de leur charge est transféré au circuit à chapelet, lors de la recharge des diodes. Celle-ci s'effectue sur commande simultanée des MOS interrupteurs. Ainsi, la totalité du temps utile est pratiquement utilisée par les diodes pour intégrer la lumière.

Développant une matrice plane de 64×64 , les *Laboratoires d'Électronique et de Physique Appliquée (LEP)* ont proposé, en 1974, un dispositif dont le schéma de principe est donné figure 23. Les photodiodes fonctionnent en régime d'accumulation et l'adressage se fait par l'intermédiaire d'un réseau de connexions croisées. Le pas des cellules est de $100 \mu\text{m}$ et le circuit complet, avec un premier étage de décodage, couvre une aire d'environ $0,8 \text{ cm}^2$.

Le *LEP* proposait d'ailleurs une technologie du type RMOS, donc recourant à des matériaux réfractaires qui utilise pour la délimitation des sources et des drains un masque de molybdène ou de tungstène dont on conserve ultérieurement les parties qui correspondent aux gates des MOS et au premier niveau d'interconnexions.

9. LE RÉSEAU A INJECTION DE CHARGES DE GE

C'est dans cette même technologie RMOS, d'ailleurs, que *General Electric* avait réalisé, dès 1971, un premier réseau de 32×32 points à « injection de charges »

(ou CID, de *charge injected device*), puis de 100×100 monté dans une caméra miniature exposée au *Salon international des Composants électroniques*, à Paris, en 1973.

L'âme de la caméra est constituée par un total de 10 000 paires de condensateurs de type MOS distribués en un réseau de 100 rangées par 100 colonnes sur une puce de silicium.

L'image lumineuse qui frappe ce réseau provoque, sous l'impact des photons, la naissance de paires électron-trou. Les charges minoritaires sont alors récupérées dans la zone d'inversion située à l'interface silicium à oxyde de silicium, sous l'électrode du condensateur MOS intéressé (fig. 24).

La lecture, elle, s'effectue non pas par déplacement de charges selon un principe d'auto-balayage, mais en injectant dans le substrat ces charges. On mesure alors tout simplement le courant de substrat qui en résulte et qui témoigne de l'éclairement reçu.

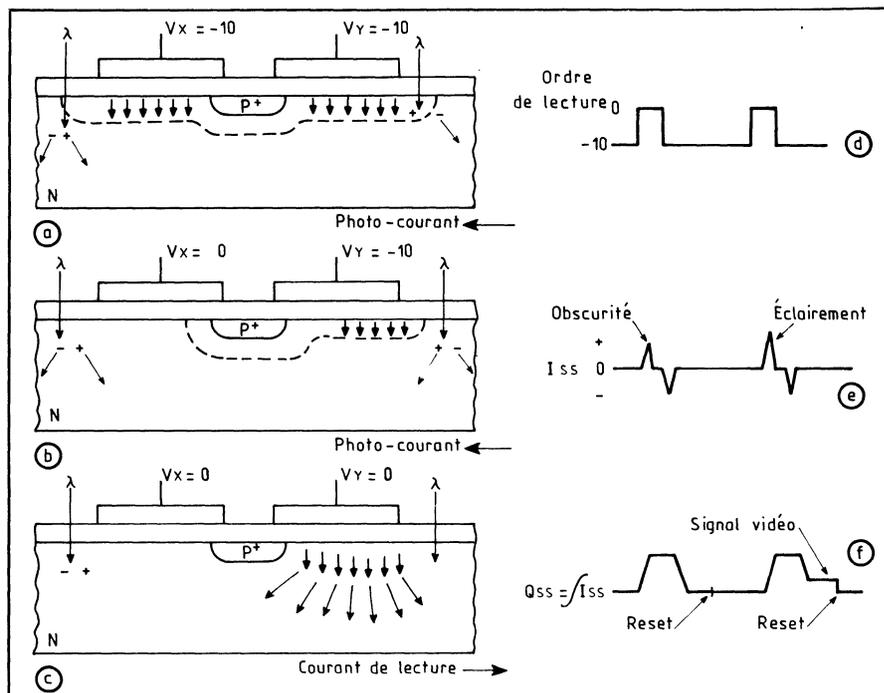


Fig. 24. — Deux condensateurs MOS reliés par une diffusion P constituent le point sensible fondamental du système à injection de charges de GE. En (a), des tensions négatives, en x et en y , sont appliquées aux deux condensateurs. En (b), l'une de ces tensions étant supprimée, les charges positives s'accumulent sous un seul élément. En (c), la polarisation ayant disparu, un courant de lecture apparaît dans le substrat. En d, allure des signaux de commande de lecture du réseau, pour un axe de coordonnées. Deux impulsions égales mais opposées sont provoquées par l'ordre de lecture si le point lu était dans l'obscurité (e); elles sont d'amplitude différente dans le cas contraire. En (f) : pour témoigner de l'éclairement et annuler ces impulsions parasites, il suffit de les intégrer.

Tel est le principe du circuit intégré à injection de charges, tel qu'il était d'ailleurs présenté (pour la première fois) par MM. G. J. Michon et H. K. Burke, de *General Electric*, à la *Conférence internationale sur les circuits à l'état solide* qui s'est tenue le 15 février 1973, à Philadelphie.

D'après le schéma, on voit (en coupe) que deux capacités MOS sont couplées par une zone diffusée P. Cet ensemble constitue le point sensible fondamental. En a une tension de -10 V est appliquée aux deux électrodes supérieures des condensateurs; les porteurs négatifs dans le silicium sous-jacent sont alors chassés. Si l'on supprime la polarisation d'une des capacités, pendant le balayage de lecture des autres points du réseau par exemple, la charge accumulée sous son électrode supérieure sera transférée à la capacité voisine via la zone P : c'est ce que montre le schéma b . Ces porteurs ne seront injectés dans le substrat que lorsque aucune polarisation ne sera appliquée aux capacités (en c).

De cette façon, la scrutation des points sensibles peut être aussi bien aléatoire que séquentielle, et commandée selon les deux axes de coordonnées. Cette dernière formule ayant été retenue, des registres à décalage MOS ont été intégrés sur la même puce de silicium sur deux côtés perpendiculaires du réseau.

Le signal vidéo brut est représenté par les charges injectées dans le substrat et est extrait lors de la lecture de chaque point sensible. Il est cependant délicat à mettre en évidence en raison des perturbations introduites lorsqu'on applique la tension de lecture : en effet, comme le montrent les courbes d et e , chaque impulsion positive donne lieu à l'apparition de deux impulsions opposées, dues aux couplages capacitifs parasites sur la ligne lue.

Celles-ci s'annuleront cependant si l'on intègre le signal de lecture (en f), car alors, on retrouve le signal de commande et non plus simplement un courant de substrat du type Cdv/dt .

Comme on le voit sur la figure, les impulsions correspondant à un point resté dans l'obscurité et à un point soumis à l'éclairement sont alors de niveau différent, proportionnel à l'éclairement.

La tension qui apparaît à la sortie du capteur est ainsi fonction de la charge totale injectée, divisée par la capacité totale de la ligne. Si celle-ci reste suffisamment faible, c'est le courant d'obscurité qui limite la sensibilité globale.

Ainsi, on constate que le balayage ne s'effectue plus comme dans les dispositifs « classiques » à transfert de charge mais exige un réseau spécial de commande qui scrute chaque point séparément. Le circuit intégré résultant semble plus complexe mais, selon ses promoteurs, il offrirait deux avantages décisifs sur les CCD :

- il est plus facile à produire industriellement;
- un défaut sur un point du réseau met hors service ce seul point et non plus toute la ligne des éléments en cause. Le point affecté procurera une tache noire, tout simplement, lors de la visualisation.

Le rapport signal sur bruit du dispositif à injection de charge est suffisant pour permettre la reproduction de 10 valeurs différentes de gris. La sensibilité permet la prise de vues à l'éclairage ambiant. Elle s'effectuait d'ailleurs à la cadence normale de 25 images par seconde.

Le réseau à injection de charge, dans sa structure à 32×32 points d'analyse, mesure moins de 1 cm^2 (80 mm^2 environ).

10. LA LOGIQUE A CCD

En raison de sa topologie très particulière, la logique à CCD va être présentée dans ce chapitre, donc avant l'analyse générale des fonctions logiques. Le lecteur qui ne connaîtrait pas le principe de ces fonctions pourra donc se reporter au chapitre qui leur est consacré avant de poursuivre. Ajoutons que la logique à CCD soulevait bien des discussions, en 1978, et n'avait toujours pas débouché commercialement.

Pour créer des fonctions logiques avec des CCD, on dispose judicieusement des électrodes de commandes sur le réseau. Selon le cas, on réalisera des fonctions *destructives* (les informations traitées sont détruites et seul subsiste le résultat) ou *non destructives*.

La figure 25 montre une porte AND : le contenu du puits placé sous l'électrode de commande C est égal à AB. Le montage est vu de dessus, c'est-à-dire du côté électrodes de commande; si la source de porteurs (une jonction PN polarisée en direct) est excitée, les charges n'atteindront C que si les électrodes A et B permettent le transfert.

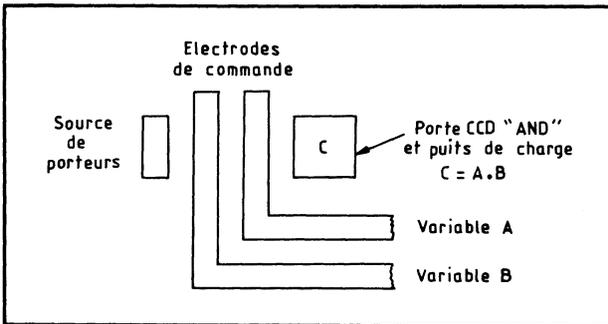


Fig. 25. — Principe de la porte AND en CCD.

Le OR répond aux mêmes principes. La source (fig. 26) est répartie en deux branches commandées chacune par une électrode métallique, A et B. Par conséquent, $C = A + B$. Dans ces deux cas, les charges ne disparaissent pas; la lecture est non destructive ou plutôt, le résultat servira généralement à une commande logique suivante.

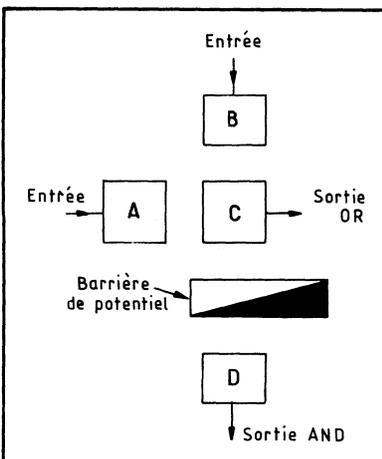
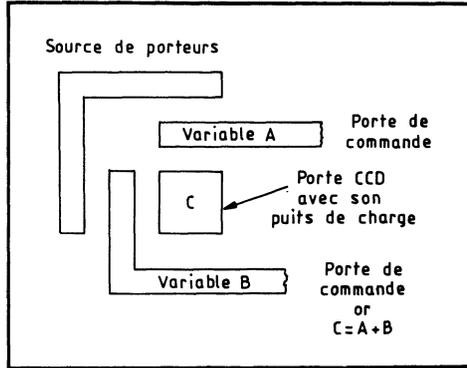


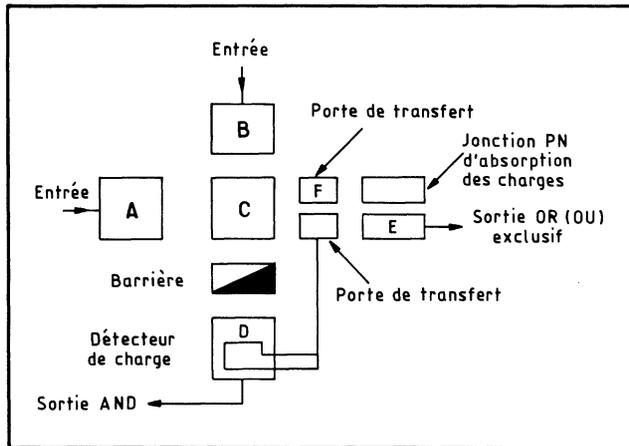
Fig. 26. — Réalisation du OR.

Fig. 27. — Fonctions complexes en CCD, fournissant $C = A + B$ et $D = AB$.



Soit, par exemple, la fonction de la figure 27. Les électrodes A et B de registres à décalage propulsent les charges qu'elles commandent sous l'électrode C. Celle-ci communique avec D, mais via une barrière de potentiel (créée par implantation ionique ou assurée à l'aide d'une tension séparée). Les « capacités » de charge de A, B, C et D sont identiques (le contenu de leur puits respectif); par conséquent, si A et B sont pleins, C débordera et c'est cet excès de charge qui passera en D, traversant la barrière de potentiel. Les fonctions réalisées sont $C = A + B$ et $D = A \cdot B$.

Fig. 28. — Constitution d'un OU exclusif.



Le OU exclusif répond au schéma de la figure 28, développement du précédent. Ici, la porte D est dotée d'un détecteur de charge. Si aucune charge n'est présente, il autorise la porte de transfert à fonctionner et des charges de C à passer en E. Par conséquent, la sortie en E est un OU exclusif. Le rôle de la porte F consiste à vider les charges piégées sous C si celles-ci n'ont pas été transférées sous E. Ce montage est simplement un demi-additionneur, additionnant A et B et fournissant la somme (E) et la retenue (D).

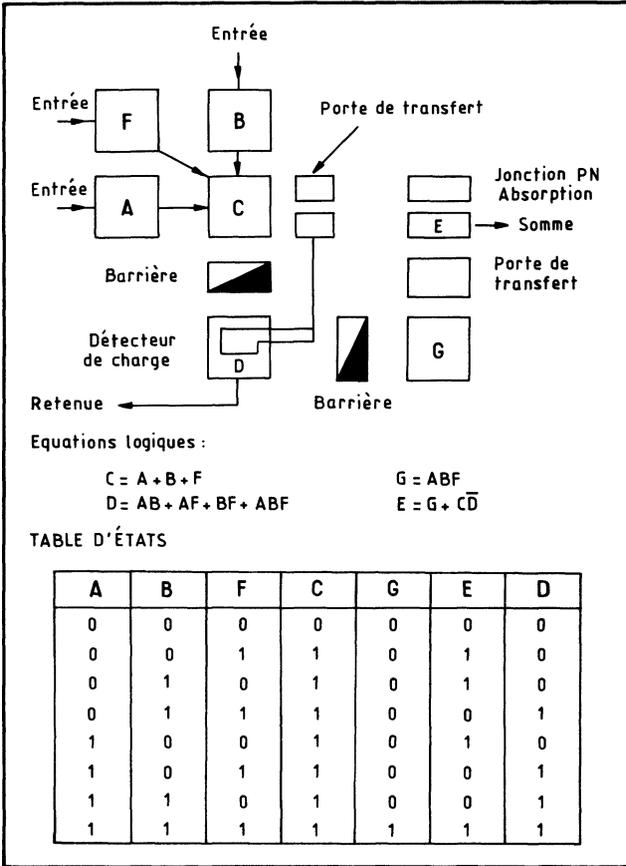
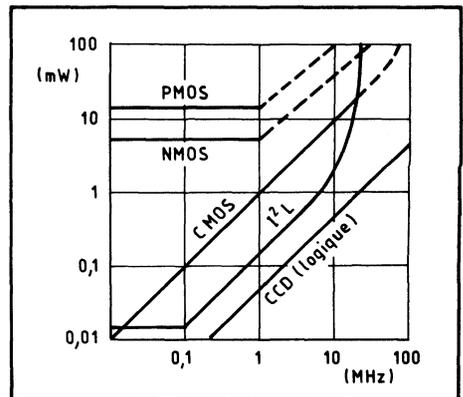


Fig. 29. — Le plein additionneur en CCD.

Fig. 30. — Variations comparées de la consommation en fonction de la technologie, avec extension aux 100 MHz des CCD, annoncées par TRW.



Réseaux logiques comparés, en projection 1980 et au-delà

Technologie	Porte logique		Etage de registre à décalage			Principe de fonctionnement
	Facteur de qualité (pJ)	Surface occupée (μm^2)	Dissipation par bit à 1 MHz (μW)	Surface occupée (μm^2)	Vitesse max. (MHz)	
CCD	0,2	1200	1	300	100	Dynamique
VLSI triple diffusion	< 5	6000	100	20000	300	Statique
I ² L	1	6000	50	20000	30	Statique
CMOS/SOS	< 2	10000	100	50000	80	Statique
Bulles magnétiques(?)	< 1	< 1000	< 1	100	1	Dynamique (non volatil)

De là, on peut passer à un plein additionneur (fig. 29), avec une 3^e entrée pour la retenue précédente, une sortie somme (le OU exclusif) et une sortie retenue suivante (le AND) : on part du fait que le puits sous D ne sera chargé que si deux au moins des trois entrées, A, B et F contiennent des charges.

En poursuivant selon des raisonnements identiques, on peut développer toutes sortes de schémas et de fonctions logiques.

En logique, l'intérêt des CCD réside dans leur grande densité d'intégration. De ce fait, on peut aisément concevoir des systèmes parallèles ou pipe-line (les opérations sont exécutées en série le long d'une chaîne dont chaque élément se voit confier « sa » fonction, au passage) qui pallient la lenteur relative inhérente à cette technologie; le gigahertz devient pensable. En outre, la programmation des fonctions peut être assurée simplement, au point qu'on pourrait imaginer un microprocesseur à CCD, ou encore un ensemble de calcul de la transformée rapide de Fourier (ou FFT : « fast Fourier transform »), si l'on en croit A. ZIMMERMAN, de TRW, et F. BARBE, du Naval Research Lab, de Washington (1). Aboutissant aux mêmes conclusions, G. DORIATH, de la division Equipements avioniques et spatiaux de Thomson-CSF proposait en 1977 un bloc de calcul à CCD pour le traitement numérique du signal (2).

Le tableau compare quatre technologies, dont les CCD, dans leurs applications aux portes logiques et aux registres à décalage. La figure 30 montre comment varie la consommation en fonction de la fréquence.

On peut imaginer, d'ailleurs, que les circuits à bulles magnétiques dont la première commercialisation date de 1977 sous forme de mémoire pourraient mener, eux aussi, à des systèmes logiques selon des organisations transposées de ce type, ce qui a été évoqué dans le tableau.

1. Communications présentées par A. ZIMMERMAN à la National Telecommunication Conference à San Diego en 1974 et San Francisco en 1975; article de D. F. BARBE et W. D. BAKERT dans *Microelectronics*, vol. 7, n° 2, 1975. On en retrouvera des éléments dans un article publié par *Electronics*, du 31 mars 1977, p. 97, et dans *Electronique et Applications Industrielles*, n° 255 du 15-6-1978.

2. *Revue technique Thomson-CSF*, vol. 9, n° 3, septembre 1977.

CHAPITRE VIII

LES VMOS

1. PRÉSENTATION DES VMOS

La plus récente des technologies appliquées aux MOS est la VMOS, de « *Vertical MOS* ». Elle confère aux transistors, soit un encombrement moindre (surface moindre de silicium par transistor), soit une puissance accrue. Les premiers VMOS commercialisés, sous forme de transistors discrets, datent de 1976 (société *Siliconix*); dès 1977, la société *AMI* les intégrait pour proposer des mémoires RAM de très hautes performances.

La figure 1 montre, en coupe, la structure d'un VMOS comparée à celle d'un MOS classique. A surface occupée égale et grâce à cette structure, les courants admis atteignent plusieurs ampères et les puissances, plusieurs dizaines de watts; la tension de claquage, elle, atteint la centaine de volts ou davantage (200 ou même 400 V), alors que la résistance du canal à l'état passant R_{on} est réduite : on trouvera, par exemple, $1,8 \Omega$ à 1 A pour tel transistor VMOS discret en boîtier TO-3. Si l'on réduit la surface occupée, on réduit courants et tensions mais on bénéficie d'une vitesse de travail accrue.

2. VMOS EN COMPOSANTS DISCRETS

Pour réaliser un VMOS, on commence par diffuser les zones de canal et de source. Puis, une gravure en V est effectuée par attaque chimique, selon des dimensions précises. Le transistor est achevé par croissance d'oxyde et métallisation.

Comme on peut en juger d'après la figure, *le courant se propage verticalement* dans le VMOS, et non plus horizontalement, de la source située au sommet de l'élément vers le drain, sur la face opposée. Le courant traverse donc les couches diffusées, au lieu de les « longer »; d'autre part, les diffusions P et N définissent la longueur du canal, qui peut ainsi être réduite à 1,5 microns (comme avec les MOS à double diffusion, ou DMOS), ce que l'on comparera aux 6 microns des

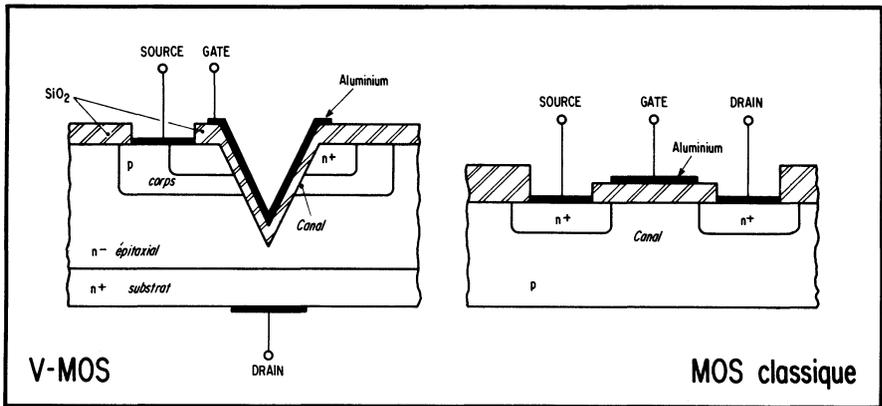


Fig. 1. — Le VMOS, comparé à un MOS courant (structure vue en coupe).

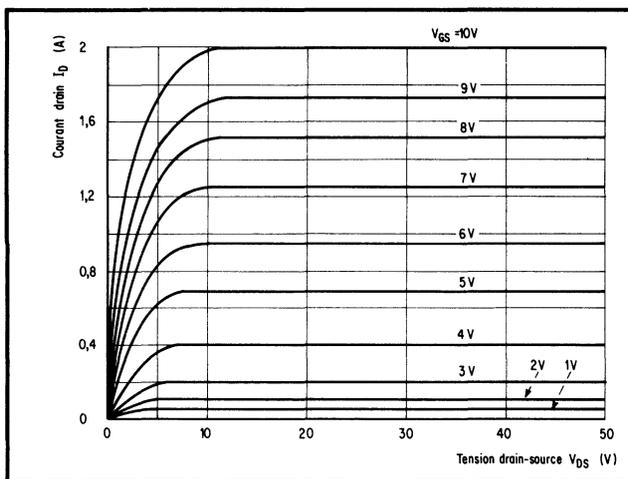


Fig. 2. — Caractéristiques courant-tension du VMOS.

MOS courants. Enfin, les deux faces du sillon en V sont conductrices. Toutes ces raisons concourent à accroître le courant et même sa densité. L'accroissement de la tension de claquage, de même que la faible capacité de réaction du VMOS, proviennent de la zone diffusée supplémentaire N⁻. En effet, sa faible concentration en impuretés amène le courant à s'étaler en accédant au drain, ce qui réduit le champ électrique et par conséquent, permet une importante augmentation des tensions utiles. D'autre part, la capacitance gate-drain est réduite précisément par la présence de la zone N⁻.

On constate encore une meilleure linéarité de la relation courant de drain fonction de la tension de gate (fig. 2). A ces deux caractéristiques, courant et tension accrus, il faut en ajouter une troisième qui est une fréquence de travail supérieure, résultant directement de la faible longueur du canal.

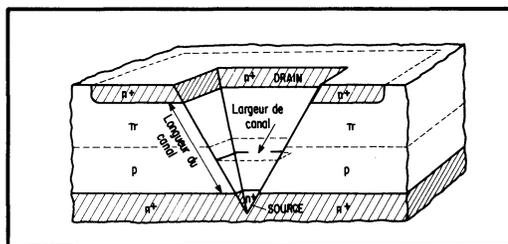
Un VMOS tel que le VMP-1, l'un des premiers commercialisés par *Siliconix*, le créateur de ces éléments, dissipe 35 W à 25 °C; avec 1 A et sous 10 V de tension gate-source, la résistance drain-source est inférieure à 2,5 ohms; sa capacité d'entrée est de 35 pF et sa capacité inverse de 8 pF; son temps de réponse pour 1 A est de 5 ns; sa résistance thermique est de 3,5 °C/W.

3. CIRCUITS INTÉGRÉS VMOS

Les premiers circuits intégrés relevant des VMOS ont été commercialisés en 1978. Par rapport à la technologie bipolaire I²L, les opérations sont résumées dans le tableau A. Le nombre d'opérations est identique.

Ces premiers circuits intégrés ont été des mémoires RAM. Ainsi, la S 4015-3 de *AMI* est une RAM statique de 1 024 mots de 1 bit, concurrente directe des célèbres bipolaires 93415/25 de *Fairchild* avec ses 50 ns de temps d'accès, mais moins chère. La version rapide de la 4015 a un temps d'accès de 30 ns.

Fig. 3. — Vue en coupe d'un VMOS intégré, selon AMI.



Dans la structure de *AMI*, à VMOS canal N, le substrat N⁺, fortement dopé, sert de source (fig. 3) et de masse commune. Le substrat effectif des MOS est une couche P, créée entre la source et les drains N⁺. La couche notée Pi, légèrement dopée P, est épitaxiale et sert de région à charge d'espace destinée à réduire la capacitance et accroître la tension de claquage de la jonction drain-substrat.

La longueur effective du canal est ramenée à l'épaisseur de la couche P, soit 1 micron environ, mais comme ce canal ceinture le sillon en V, sa « largeur » est importante. Pour commander un courant identique à celui de MOS classiques, on peut donc réduire fortement toutes les dimensions superficielles.

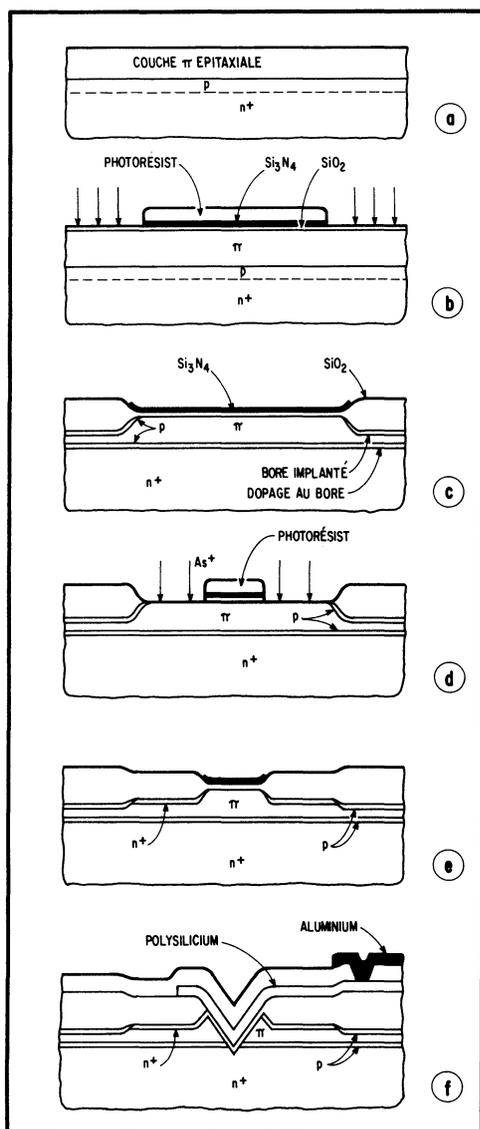


Fig. 4. — Film de la fabrication d'un VMOS intégré, commenté dans le texte.

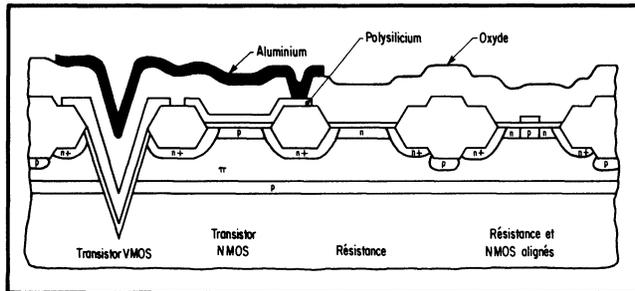
De ce fait, la RAM bipolaire 93415 est 60 % plus vaste, en surface, que la 4015 en VMOS. La surface de cette dernière pourrait encore être réduite de moitié! Ainsi, une RAM de 1 K pourrait n'occuper que 4 500 mils², soit environ 3 mm², tout en travaillant à 28 ns de temps d'accès. On notera que le substrat commun N^+ , qui sert de source commune pour tous les VMOS, contribue fortement à cette réduction de surface en économisant les lignes habituelles de masse. Les traits de la première RAM en VMOS restaient au pas courant de 5,5 microns.

A l'avenir, des temps d'accès de 20 ns sont concevables. Les densités s'accroîtront également; ainsi, une ROM de 65 536 bits pourrait tenir sur 28 000 mils²,

soit environ 18 mm^2 . Les premiers microprocesseurs en VMOS étaient également attendus pour 1978-1979. Dès 1977, le facteur de qualité des VMOS était de $2,5 \text{ pJ}$, pour une consommation de $0,05 \text{ mW/bit}$ et une vitesse de 50 ns , l'alimentation unique étant de $+5 \text{ V}$.

Le processus de fabrication d'un VMOS intégré est développé figure 4 (selon AMI); il demande 7 masquages. On commence par doper la surface d'une tranche, fortement dopée N^+ , avec du bore, puis on fait croître la couche II épitaxiale faiblement dopée P, sur 3 microns environ (en *a*). On dépose ensuite de la silice et du nitrure de silicium, ce dernier étant gravé pour préparer une implantation ionique de bore (b). Ainsi, deux fines couches P ont été formées. On fait alors

Fig. 5. — Structure d'un ensemble intégré VMOS, résistance et NMOS.



croître de l'oxyde (*c*) et l'on grave à nouveau le nitrure pour préparer un dopage N^+ de drains, à l'arsenic, assuré également par implantation ionique (en *d*). Une oxydation épaisse servira alors à recouvrir les zones diffusées (*e*). Les sillons en V sont enfin gravés, on fait croître l'oxyde de gate, on dépose du polysilicium (silicium polycristallin), les contacts sont gravés et l'on termine par les métallisations (*f*).

Cet unique processus peut mener à la fabrication de trois types de composants (fig. 5) : le VMOS décrit, mais aussi un NMOS ou une résistance, ces deux derniers pouvant d'ailleurs être alignés.

Afin de mieux situer la technologie VMOS, le tableau B caractérise quelques types de circuits mémoires réalisés selon des technologies concurrentes qui sont la HMOS, la VMOS, la I2L en version Schottky, la I3L (pour I2L par implantation ionique), et la TTL Schottky qui sert bien souvent de référence, avec ici la 93415, mémoire de Fairchild.

En 1978, la technologie VMOS s'opposait ainsi à la HMOS, ou MOS à Hautes performances imaginée par Intel. Dans la HMOS, toutes les dimensions des transistors ont été réduites selon une même échelle, sans aucun changement de technologie de base; ainsi, le canal de $6 \mu\text{m}$ est passé à $4 \mu\text{m}$ en un premier temps, le reste à l'avenant. De ce fait, surface occupée et consommation diminuent alors que la densité et la vitesse s'accroissent. Mais une telle formule pourrait, elle aussi, s'appliquer aux VMOS auxquels elle ajouterait ses qualités.

4. LES VMOS DISCRETS EN COMMUTATION

Les VMOS discrets, en raison de la puissance importante qu'ils peuvent piloter à partir d'une puissance de commande infime, sans comparaison avec

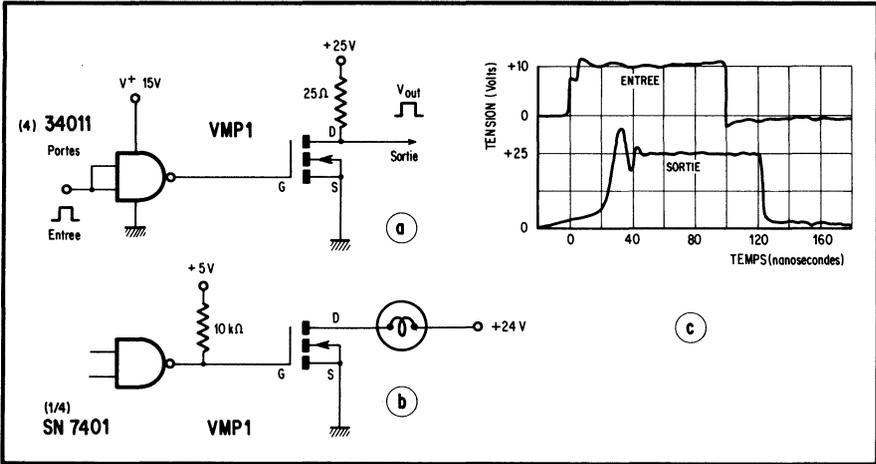


Fig. 6. — Commande d'un VMOS par des portes CMOS (a) et réponse-type (b); commande par une porte bipolaire (c).

celle qu'exigent les transistors bipolaires, constituent d'excellentes interfaces entre la logique ou les microprocesseurs et les applications (périphériques ou systèmes industriels).

Par exemple, le VMP-1 cité peut, en pratique, commuter 1 A entre 1 et 4 nanosecondes, ce qui signifie qu'il est quelque 100 fois plus rapide qu'un Darlington bipolaire. Sa résistance d'entrée est celle des MOS : environ 1 000 mégohms, et de ce fait il se contente d'une commande à faible puissance.

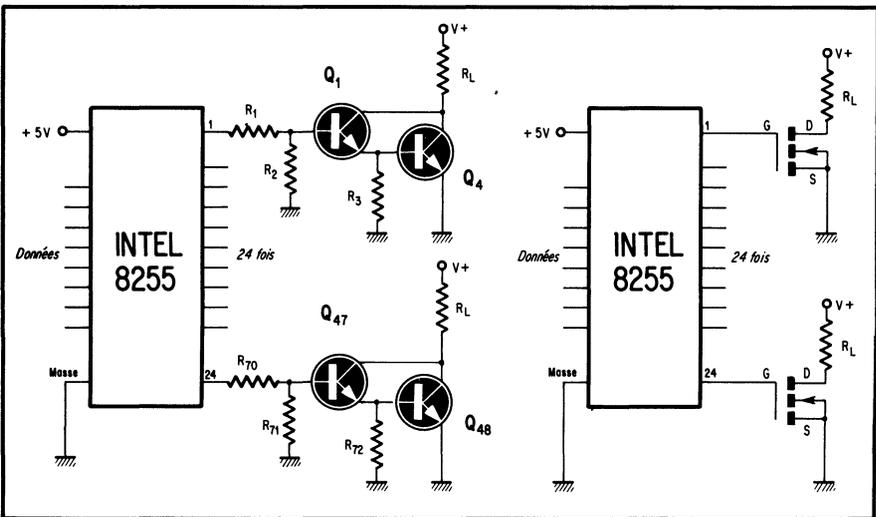
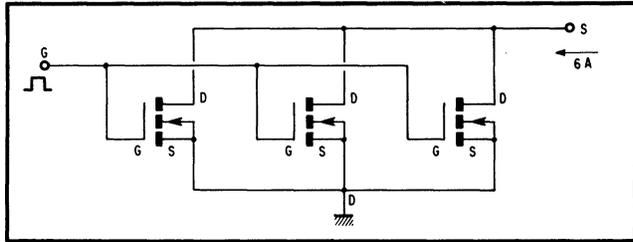


Fig. 7. — Dans un système à microprocesseur 8080, l'usage de VMOS peut simplifier considérablement l'interfaçage avec des imprimantes, solénoïdes et autres actuateurs.

De plus, et parce qu'il s'agit d'un MOS à enrichissement, une tension d'entrée nulle le bloque; son courant de fuite est alors inférieur à 0,5 microampères.

Avec une tension d'entrée de 10 V et par conséquent un courant de quelques microampères destiné à charger sous cette tension la capacité d'entrée (35 à 40 pF) en moins d'une microseconde, un courant de 1 A traversera le transistor qui présentera une résistance drain-source de 20 ohms. Pour obtenir le même résultat avec un transistor bipolaire, il aurait fallu 220 mA de courant de commande.

Fig. 8. — Montage parallèle direct de VMOS.



Le schéma d'un VMP-1 commandé directement par 4 portes CMOS en parallèle est donné figure 6a. La commutation de 1 A est assurée en 20 ns, les ondes étant dessinées en b. Avec une seule porte, le temps de commutation passerait à 50 ns. Une porte TTL à collecteur ouvert (c) sera chargée par 10 kΩ ou moins pour accroître la vitesse (ou davantage pour réduire la consommation!).

En matière de commutation de courant *Siliconix* propose une autre application relevant de la micro-informatique (fig. 7). Le circuit d'interface 8255 du système microprocesseur 8080 de *Intel* peut être suivi, pour commander des imprimantes, solénoïdes ou autres actionneurs, soit par 48 transistors bipolaires plus 72 résistances, soit par 24 VMOS. Dans ce dernier cas, de surcroît, la consommation est ramenée de 240 mW sous 5 V à 1,2 μW (10 nA de fuite par VMOS), toujours sous 5 V.

Pour accroître encore le courant, on monte directement en parallèle les VMOS (fig. 8). Le coefficient positif de température de la résistance drain-source tend à limiter et par conséquent à égaliser les courants, à l'inverse de ce qui se passe dans les bipolaires. Les 3 éléments 2 A de la figure fournissent ainsi 6 A.

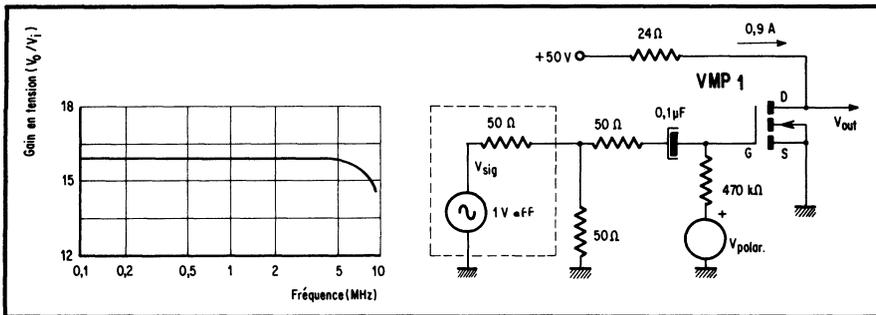


Fig. 9. — VMOS en commutation analogique.

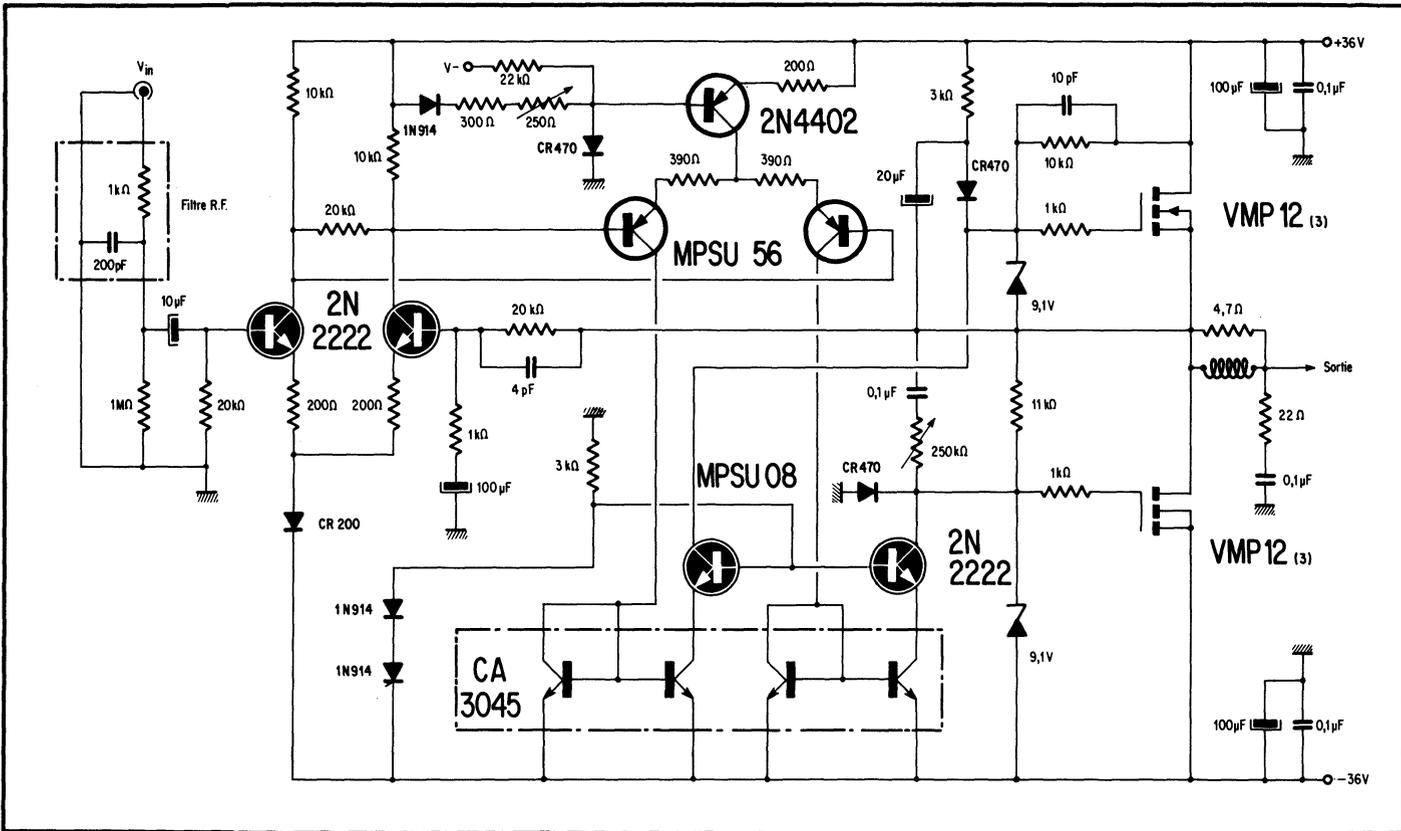


Fig. 10. — Amplificateur hi-fi 80 W (un canal).

5. LES VMOS DISCRETS EN LINÉAIRE

Puisqu'on en est aux composants discrets et que ces VMOS constituent une technologie qui n'est pas encore entrée dans les applications courantes, on va en

Tableau A. — Processus comparés pour circuits intégrés

VMOS	I ² L
1- Diffusion de source N	1- Diffusion d'émetteurs N
2- Croissance épitaxiale P	2- Epitaxie N
3- Diffusion de drain N	3- Bases et injecteurs P
4- Gravure profonde	4- Gravure de collecteur
5- Réalisation des gates	5- Formation Schottky
6- Contats	6- Contacts
7- Métallisation	7- Métallisation

Tableau B. — Comparaison entre des mémoires réalisées selon diverses technologies

Technologie	1978-1979			Projection au-delà de 1980		
	Accès (ns)	Cycle (ns)	Consommation (mW) Activité/Veilleuse	Densité d'intégration (Kbits en mémoires)	Tension d'alimentation (V)	Technologie éprouvée (en 1978)
HMOS (statique)	50	50	500/50	≥ 64	3 (?)	Oui
VMOS (statique)	50	50	500/50	≥ 64	5	Trop jeune
I ² L Schottky (statique)	70	700	400/40	≥ 64	(5)	Oui
I ³ L (dynamique)	100	200	500	≥ 64	(5)	Oui
TTL Schottky (statique)	30	30	750	< 16	5	Oui

illustrer rapidement ici quelques applications en linéaire. En effet, le gain en courant d'un transistor VMOS est équivalent à celui d'un transistor bipolaire à bêta infini; le produit gain-bande unitaire se manifeste à 600 MHz (et ce, en tenant toujours compte des premiers VMOS); enfin, la caractéristique du courant drain-source est linéaire de 400 mA à 2 A.

La figure 9 montre ainsi le VMP1 en commutateur analogique à faible résistance (1,5 à 3 Ω), faible fuite (0,5 μA), transconductance de 0,27 mhos, et gain de 6,5 avec charge de 24 Ω. La bande passante va du continu à 10 MHz et la distorsion harmonique est de 0,075 % à 1 V eff de sortie, de 0,8 % à 10 V eff.

Tableau C. — Quelques technologies évoluées

Technologie	Densité (portes/mm ²)	Temps de propagation (ns)	Facteur de qualité (pJ)
HMOS 1978	170	1	1
HMOS >1980	220	0,4	0,2
VMOS 1978	220	1	1
SOS 1978	150	0,5	0,2
I2L 1978	120	10	1
I2L >1980	200	3	0,2

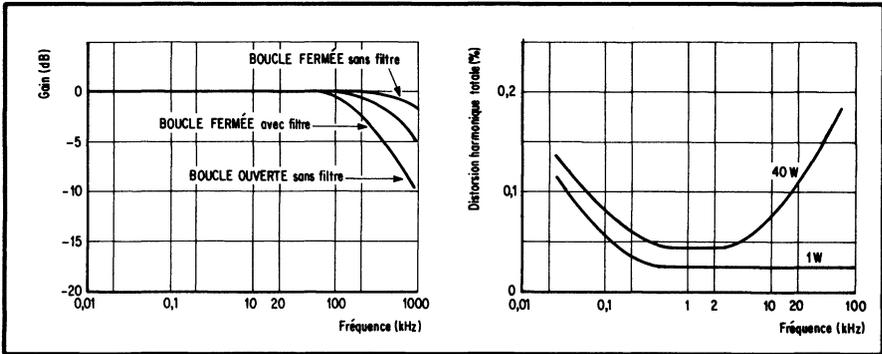


Fig. 11. — Courbes de réponse de l'amplificateur de la figure précédente.

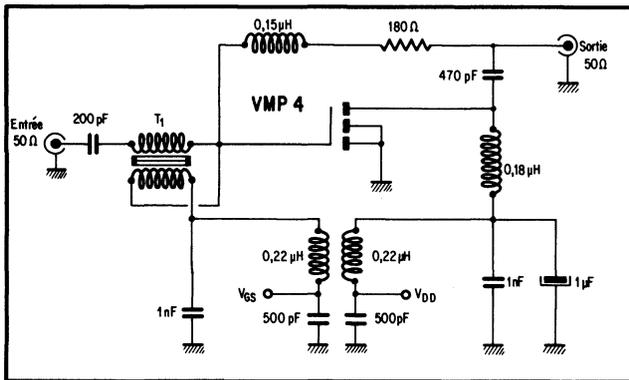


Fig. 12. — Amplificateur HF de 40 à 180 MHz.

De là à passer à un amplificateur hi-fi... C'est ce que montre la figure 10, toujours selon les études de *Siliconix*, pour un amplificateur 80 W, stéréo, mais pour un seul canal où l'on utilise 2 fois 3 VMOS en parallèle. Bien que la contre-réaction soit très modeste (22 dB), la distorsion harmonique n'est que de 0,04 % pour une réponse de 3 dB de 1 Hz à 800 kHz (fig. 11).

En haute fréquence, le schéma de la figure 12 fait intervenir un modèle VMP-4, en montage stripline, avec un gain de puissance de 15 dB plat à 1 dB entre 40 et 180 MHz. Ce circuit débite 10 à 12 W dans 50 Ω .

L'intérêt des VMOS sera encore multiplié lorsque des éléments complémentaires, canal N et canal P, seront disponibles. Ils étaient annoncés fin 1978 par divers fabricants, dont *ITT-Intermetall* avec les BD522 en canal N et BD512 en canal P, composants de 60 V sous 1,5 A maximum.

CHAPITRE IX

L'INVERSEUR FONDAMENTAL

Le circuit fondamental qui donne accès à tous les schémas est, pour les MOS, l'inverseur. Il peut être à charge passive, mais aussi et surtout à charge active, un second MOS servant alors de charge. Dans ce cas, un très grand nombre de combinaisons sont possibles. Enfin, l'inverseur fondamental peut être statique ou, et c'est une notion nouvelle qui apparaît avec les MOS, dynamique.

1. L'INVERSEUR A CHARGE PASSIVE

L'étage d'inversion le plus élémentaire est constitué par un MOS chargé par une résistance (fig. 1). L'entrée se fait sur le gate et la sortie sur le drain; le substrat est connecté à la tension la plus positive pour un PMOS (la plus négative pour un NMOS).

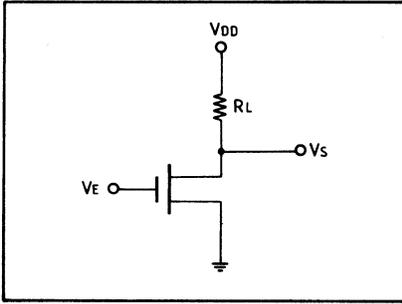


Fig. 1. — Inverseur chargé par une résistance. La tension d'entrée peut être notée V_E ou V_G (pour G gate)

Si l'on raisonne à partir du MOS de base, c'est-à-dire d'un MOS canal P, à enrichissement, la tension V_{DD} doit être négative. On recourt alors généralement à la logique négative : le 1 logique est la tension la plus négative (elle peut être V_{DD}), tandis que le 0 logique est la tension la moins négative, donc ici la masse.

En fonctionnement logique, le transistor MOS est soit *bloqué*, soit *conducteur* :

- lorsqu'il est bloqué, donc lorsque la tension d'entrée $V_G = 0$, la sortie est V_{DD} ;
- lorsqu'il est passant, donc avec $V_G = 1$, la sortie est proche de la masse.

La tension de commande V_G doit se situer franchement en deçà ou au-delà de la tension du seuil du MOS afin que la commutation soit franche.

A l'état passant, la tension réelle de sortie dépend bien évidemment de la résistance R_{on} du MOS à l'état conducteur. Si celui-ci travaille dans la région triode

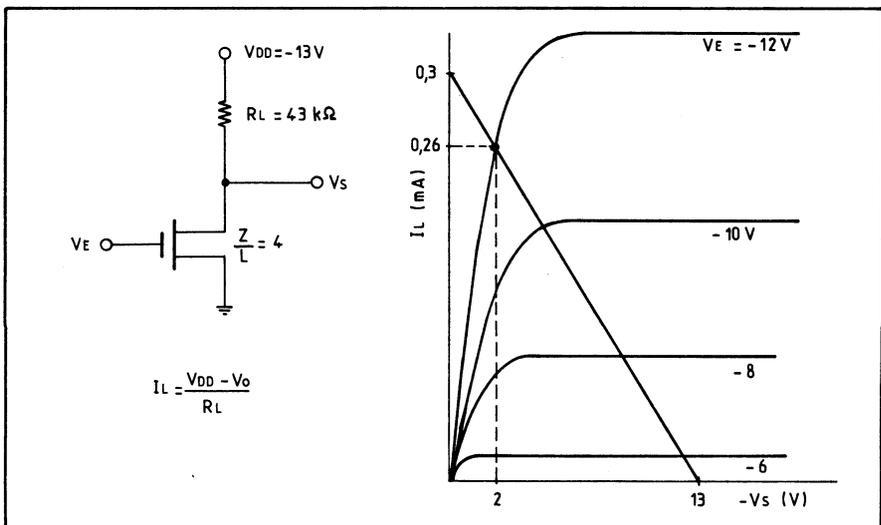


Fig. 2. — Droite de charge de l'inverseur à résistance.

et si V_{DD} est faible, la résistance $R_{DS(on)}$, notée plus simplement R_{on} est donnée, comme on l'a vu, par :

$$R_{DS(on)} = \frac{1}{K} \cdot \frac{1}{\frac{W}{L} \cdot (V_G - V_{TH})}$$

Si l'on suppose que K , constante liée à la technologie, vaut $4 \mu A/V^2$; que le rapport $W/L = 4$, et que, pour le MOS canal P à haut seuil, $V_{TH} = -4 V$ pour $V_G = -12 V$, on trouve :

$$R_{DS(on)} = \frac{1}{16 \cdot 10^{-6}} \cdot \frac{1}{8} = 7800 \Omega$$

Or, on peut supposer que la sortie de l'inverseur attaque les entrées d'autres portes identiques. Dans ce cas, il faut impérativement que la tension de sortie à l'état *bas* soit nettement inférieure à la tension du seuil, ici $4 V$, pour commander franchement les portes suivantes. Si l'on se fixe un niveau zéro, de tension maximale $V_O = 2 V$, ce qui justifie d'ailleurs l'hypothèse du fonctionnement triode, on détermine la valeur R_L de la résistance passive de charge tout simplement par :

$$\begin{aligned} R_L &= \frac{R_{on}(V_D - V_O)}{V_O} \\ &= \frac{7800 \cdot 11}{2} = 43 \text{ k}\Omega \end{aligned}$$

Le courant qui traverse la chaîne de division est donc, à l'état passant et avec un fonctionnement triode, de :

$$I_L = \frac{V_{DD}}{R_{on} + R_L} = \frac{13}{(7 + 43)10^3} = 0,26 \text{ mA}$$

On pourrait d'ailleurs écrire, ce qui revient au même :

$$\begin{aligned} I_L &= (V_{DD} - V_S)/R_L \\ &= 11/43 \cdot 10^3 = 0,26 \text{ mA} \end{aligned}$$

Dans ces conditions, la consommation en puissance P_C est de :

$$P_C = 13 \cdot 0,26 \cdot 10^{-3} = 3,4 \text{ mW}$$

La droite de charge

Il est facile de tracer la droite de charge sur le réseau donnant la tension de sortie V_S en fonction du courant I_L qui traverse la charge R_L . Les deux points-repères sont :

1. Le MOS est bloqué, donc $I_L = 0$ et $V_S = -13 V$.
2. V_{DD}/R_L , soit ici $13/43 \cdot 10^3 = 0,3 \text{ mA}$.

Cette droite de charge, tracée figure 2, coupe bien évidemment la courbe $V_G = -12 V$ à $V_O = -2 V$.

D'autres valeurs attribuées à R_L agissent directement sur le niveau *bas* de sortie (niveau 0), ce qui paraît évident puisque c'est le rapport de division en tension de l'étage qui change. On modèle ainsi les courbes de transfert (fig. 3) qui sont d'autant plus abruptes que le rapport des résistances (R_{charge}/R_{on}) est élevé.

Un tel inverseur présente un gros inconvénient : la résistance de charge est de valeur très élevée; elle est difficile à intégrer, sauf si l'on consent à lui accorder

une place importante sur le silicium au détriment de la densité d'intégration. Cet obstacle majeur a conduit à étudier d'autres possibilités, où la charge devient un élément actif, un autre MOS.

Tableau A. — Différences typiques de structures du MOS, selon fonction (MOS fondamental, pour une technologie définie)

MOS	W (μm)	L (μm)	Rapport W/L	Surface du canal (μm^2)	Ordre de grandeur de R_{on} (k Ω)
Inverseur	32	8	4	256	8
Charge	12	24	0,5	288	64
Buffer	320	8	40	2560	0,8

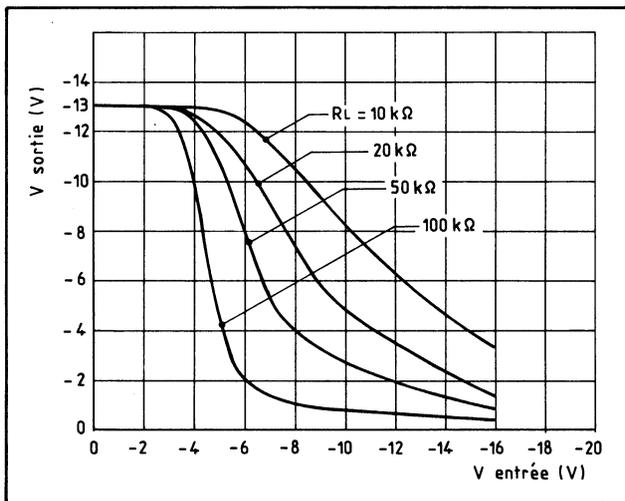


Fig. 3. — Caractéristiques de transfert pour différentes valeurs de la résistance de charge R_L .

2. LE MOS UTILISÉ COMME CHARGE

Principes

On a vu, en effet, que la résistance drain-source du MOS à l'état conducteur, R_{Dson} , ou R_{on} , dépendait du rapport W/L , pour une technologie donnée.

Dans le cas précédent, avec $W/L = 4$, on trouvait $R_{on} = 7\,800\ \Omega$. Pour obtenir une charge de $43\ \text{k}\Omega$, il suffit donc de faire $W/L = 0,7$; en effet, la même formule donne alors :

$$R_{on} = \frac{1}{4 \cdot 10^{-6} \cdot 0,7} \cdot \frac{1}{8} = \frac{10^6}{22,4} = 44\ \text{k}\Omega$$

Dès lors, un MOS doté de ce rapport peut servir de charge active, *mais avec un avantage considérable sur la résistance passive : il occupe une surface de silicium bien moindre.*

En pratique, ce qui différencie d'ailleurs un MOS en charge d'un MOS inverseur, ou même d'un MOS en buffer de puissance, c'est ce fameux rapport W/L . Des valeurs typiques sont données dans le tableau A. La dimension minimale de $8\ \mu\text{m}$ est imposée par la technologie envisagée ici.

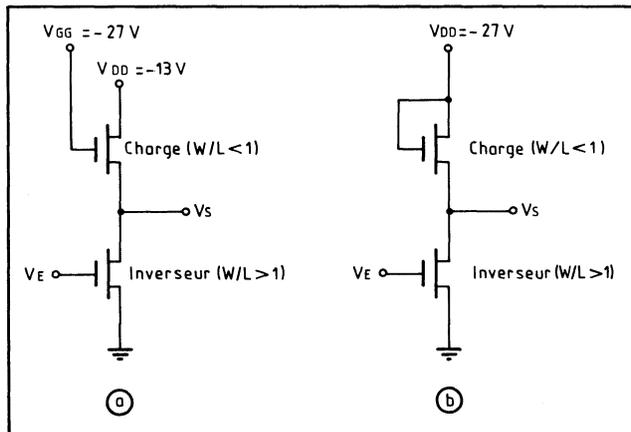
A son tour, le MOS en charge peut fonctionner en régime triode ou en régime pentode. Les différences sont les suivantes.

Fonctionnement dans la zone non saturée

Pour que le MOS en charge travaille dans la région *non saturée* (triode), son gate est porté à une tension V_{GG} supérieure (négativement) d'au moins une tension de seuil V_{TH} à V_{DD} ; on a :

$$|V_{GG} - V_{TH}| \geq V_{DD}$$

Fig. 4. — Les deux montages d'inverseurs avec MOS en charge (P, à enrichissement).



Ainsi, dans la figure 4 a, on a fait $V_{DD} = -13\ \text{V}$ et $V_{GG} = -27\ \text{V}$, pour un $V_{TH} = -4\ \text{V}$. Les tensions de sorties V_s sont :

- *Inverseur bloqué* : la sortie passe au niveau 1, proche de V_{DD} , donc de $-13\ \text{V}$.
- *Inverseur conducteur* : la sortie passe au niveau 0, proche de la masse.

Le courant I_L dans la charge est donné par :

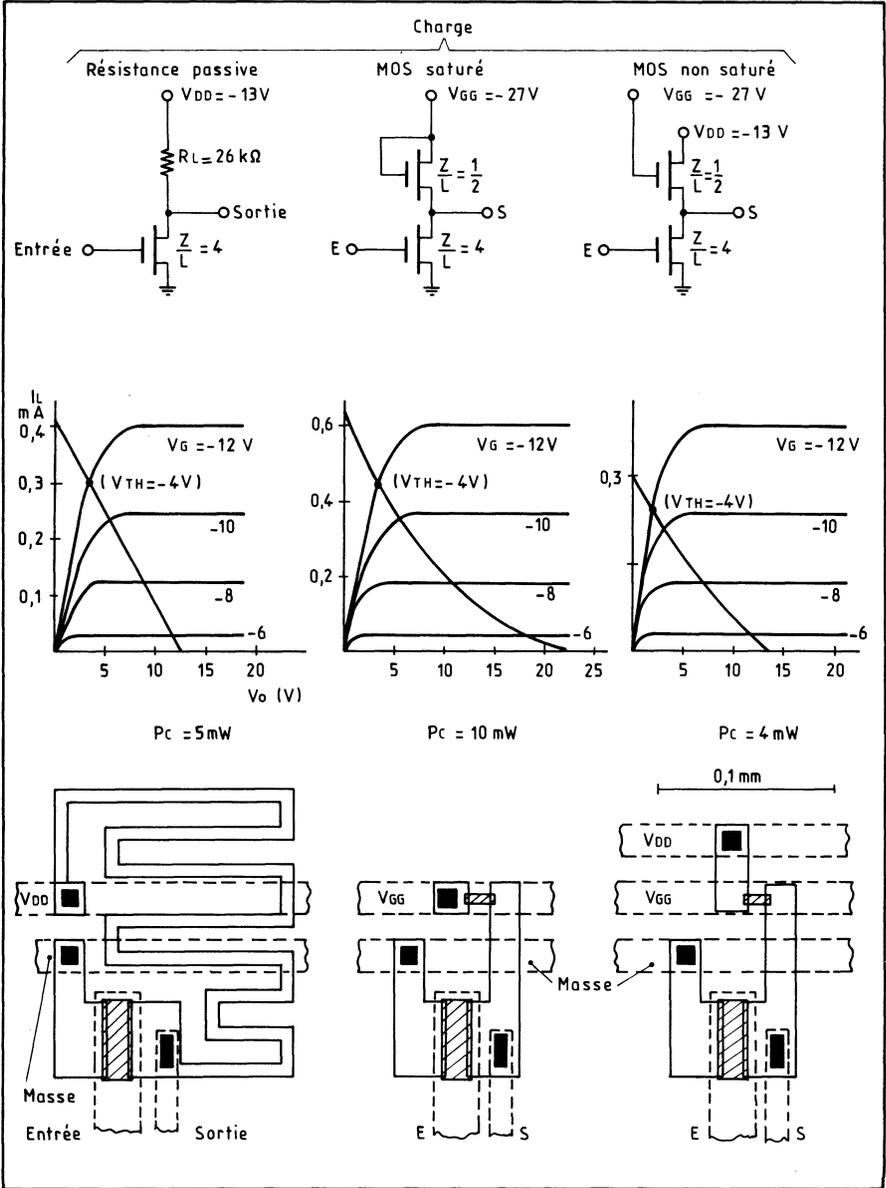


Fig. 5. — Trois montages fondamentaux d'inverseurs : schémas, droites de charge et topologies.

$$I_L = \frac{K}{2} \cdot \frac{W}{L} [2(V_{GG} - V_{DD} - V_{TH}) \\ (V_{DD} - V_S) + (V_{DD} - V_S)^2]$$

soit, avec les valeurs adoptées de $W/L = 0,7$, puis $K = 4 \mu A/V^2$ et $V_S = 2 V$:

$$I_L = 0,14 \text{ mA}$$

Fonctionnement dans la zone saturée

Pour passer en zone saturée, il suffit de relier le gate et le drain de la charge. *Le bénéfice immédiat, c'est qu'on supprime une source d'alimentation.* On fait alors $V_{DD} = -27 V$ commune (fig. 4 b). A nouveau, on trouve les deux états :

- *Inverseur bloqué* : la sortie passe à un niveau 1 proche de $V_{GG} - V_{TH}$, soit environ 22 V.

- *Inverseur conducteur* : la sortie au niveau 0 est proche de la masse.

Le courant I_L qui traverse la charge est donné par :

$$I_L = \frac{K}{2} \cdot \frac{W}{L} [(V_{DD} - V_{TH}) - V_S]^2$$

Soit, toujours avec les mêmes valeurs :

$$I_L = 0,6 \text{ mA}$$

Avantages et inconvénients

Dans les deux cas où le MOS est en charge, on note une réduction de la surface occupée, par rapport au montage où une résistance passive sert de charge à l'inverseur. C'est ce que montre la topologie tracée figure 5 (document : *SGS-Ates*), avec les courbes typiques de charge pour les trois montages retenus.

En matière de consommation, le MOS non saturé en charge est plus économe que le MOS saturé en charge : cela se conçoit, ne serait-ce que parce que la tension appliquée aux bornes du montage double d'un cas sur l'autre. La puissance maximale dissipée par *la charge seule*, si l'on reprend toujours les mêmes valeurs, sera ainsi de :

- 2,8 mW pour une résistance passive;
- 1,5 mW pour un MOS non saturé;
- 14 mW pour un MOS saturé.

En pratique, les ordres de grandeur changent, car les montages ne sont pas constamment dans une situation de dissipation maximale. De ce fait, on considère que l'étage à MOS saturé en charge consomme environ 2 fois plus que celui à MOS non saturé en charge, ce dernier absorbant lui-même une puissance à peine inférieure à celle de l'inverseur à résistance passive.

D'autre part, les fonctions de transfert sont, ici encore, d'autant plus abruptes que le rapport des résistances R_L/R_{on} , et partant W/L (*charge*) sur W/L (*inverseur*) est élevé; c'est ce que montrent les courbes des figures 6 pour une charge non

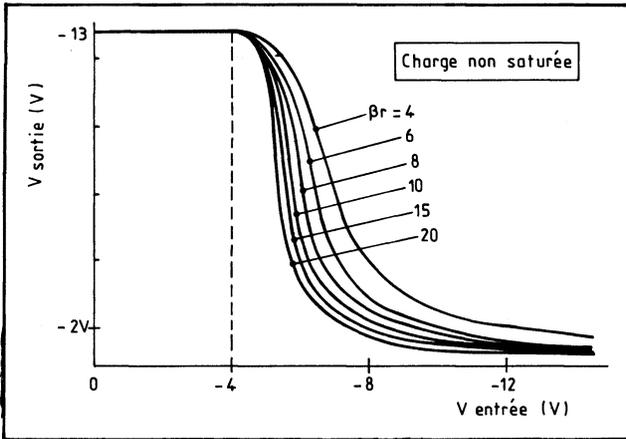
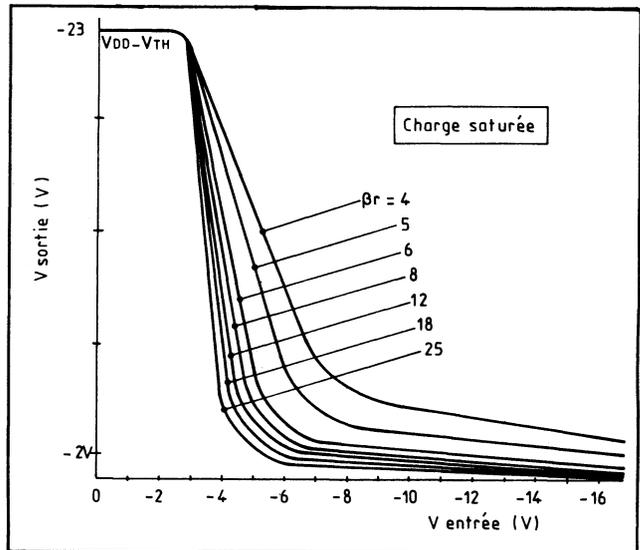


Fig. 6. — L'inverseur à charge MOS non saturée.

Fig. 7. — Inverseur à charge saturée.



saturée, et 7 pour une charge saturée. Dans ces courbes, les valeurs β_r sont les rapports β définis lors de l'étude du MOS, ce qui correspond pratiquement dans ce cas aux rapports W/L de la charge à W/L de l'inverseur; à noter que les β_r donnés ici le sont à titre indicatif et varieront beaucoup avec les technologies à bas seuil (généralement en augmentation).

3. L'INVERSEUR FONDAMENTAL EN COMMUTATION

Dans le cas le plus général, on suppose que l'inverseur est chargé par une capacité pour étudier son fonctionnement en dynamique. C'est parfaitement le cas

Fig. 8. — L'action de base de l'inverseur : charger et décharger la capacité C de la porte suivante.

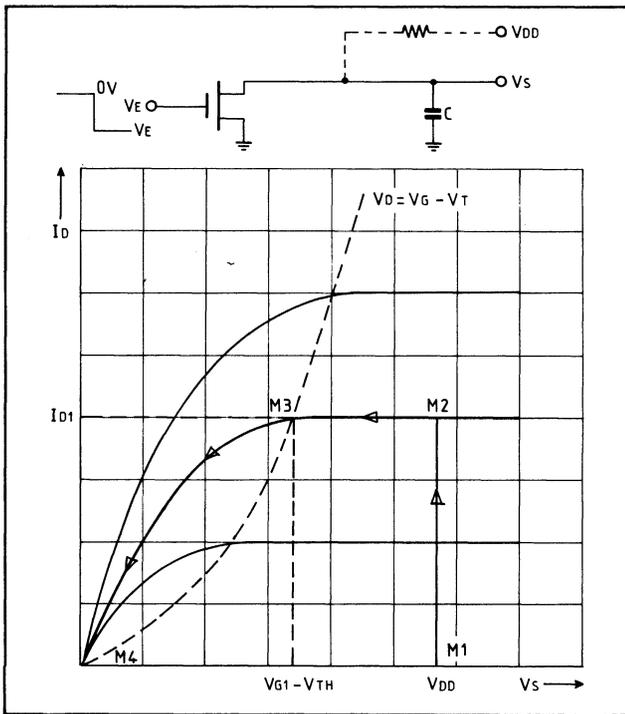
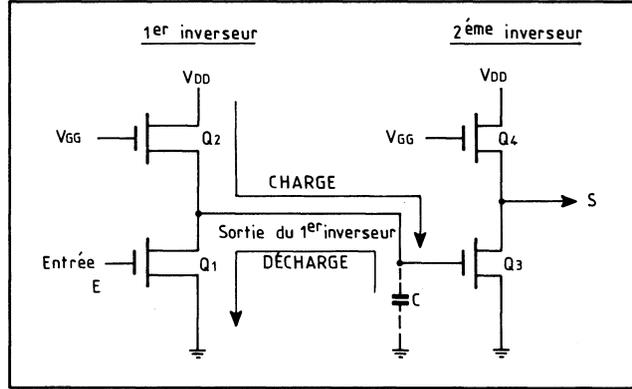


Fig. 9. — Évolution de la tension de sortie avec application d'une tension en marche à l'entrée. La tension V_D est la tension au drain du MOS alors que V_G , tension de gate, est la tension d'entrée V_E ; V_T est la tension de seuil (ou V_{TH}).

lorsqu'un inverseur est chargé par un second inverseur (fig. 8); on voit que la fonction du premier (Q_1 , Q_2) consiste, en réalité, à charger ou décharger la capacité C, qui est la capacité de structure gate-substrat du MOS noté Q_3 du second.

Mise en service de l'inverseur

Examinons ce qui se passe lorsqu'on applique un créneau de tension, de 0 V à $-V_E$, à l'entrée du premier inverseur (précédemment bloqué), en négligeant le rôle de la charge (fig. 9).

Au temps 0, la tension de sortie est V_{DD} . La tension de gate passant de 0 à $-V_E$, le point de fonctionnement caractéristique du MOS se déplace de M_1 à M_2 sur le réseau de caractéristiques I_D, V_S . Un courant I_{D1} qui décharge C traverse alors le transistor inverseur. Tant que $V_D > V_E - V_{TH}$, ce courant reste constant car le MOS est saturé; C se déchargeant, V_D diminue, et lorsque $V_D < V_E - V_{TH}$, le MOS entre dans la zone triode. Le temps total de décharge est donc la somme des temps partiels de fonctionnement en pentode et triode.

D'autre part, on notera que la charge et le MOS inverseur sont simultanément conducteurs : un chemin continu de passage du courant, de V_{DD} à masse, existe donc et il contribue à la consommation totale de l'étage.

Blocage de l'inverseur

Pour rebloquer l'inverseur, on lui applique un créneau inverse (fig. 10). Le MOS inverseur cesse de conduire et n'intervient donc plus, seule, la charge va entrer en fonction afin de fournir le courant de recharge à C .

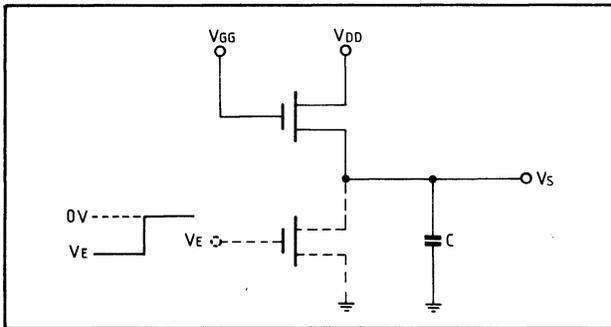


Fig. 10. — Pour une marche du niveau 1 au niveau 0, le MOS inverseur se bloque.

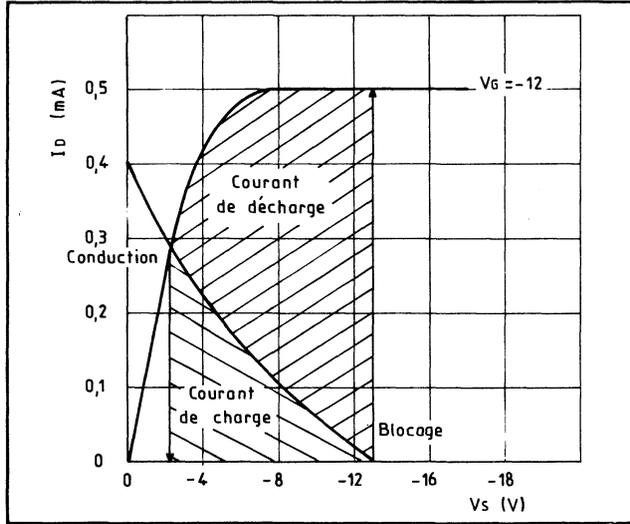
Si la charge est purement résistive (une résistance passive), le courant de charge sera une exponentielle classique, de constante de temps $T = RC$. La courbe sera un peu plus complexe avec un MOS en charge, selon son régime de fonctionnement; le temps de charge sera plus court si le MOS fonctionne en régime non saturé, avec une tension V_{GG} élevée (ou une faible tension de seuil).

En règle générale, le MOS en charge est de surface plus faible et ne peut fournir qu'un courant moindre que le MOS inverseur; en conséquence, la charge de la capacité C est plus longue que sa décharge (fig. 11). Lors de la charge de C , enfin, il n'existe pas de chemin continu V_{DD} à masse pour le courant puisque l'inverseur est bloqué, cela à l'exception des inévitables courants de fuite qu'on a systématiquement ignorés jusqu'à présent en raison de leur très faible valeur (due à la résistance élevée équivalente au MOS bloqué).

4. MOS A DÉPLÉTION EN CHARGE

Dans l'étude précédente, on a examiné le cas des MOS à enrichissement, à canal P, l'un en charge du second. Une autre formule consiste à utiliser un MOS

Fig. 11. — Caractéristiques de charge et décharge du condensateur.



à déplétion, canal P, en charge du MOS inverseur à enrichissement, canal P également (fig. 12).

Le MOS à enrichissement étant conducteur en l'absence de signal de gate, on relie son gate à la source. A nouveau, on n'a besoin que d'une seule tension d'alimentation, V_{DD} , comme lorsque la charge est un MOS saturé à enrichissement; mais à la différence avec ce dernier, la tension de sortie au niveau 1 est V_{DD} et non plus $V_{DD} - V_{TH}$, comme le montrent les fonctions de transfert de la figure 13 tracées toujours en fonction des rapports W/L , charge à inverseur.

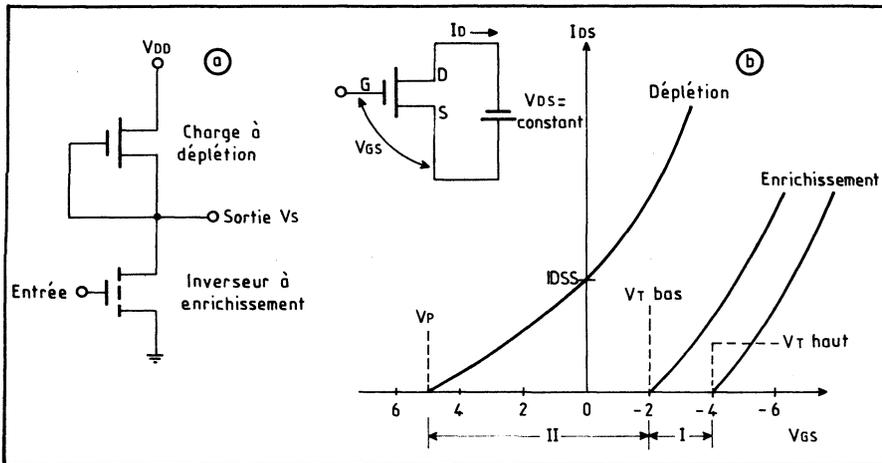


Fig. 12. — Le MOS à déplétion en charge d'un inverseur classique.

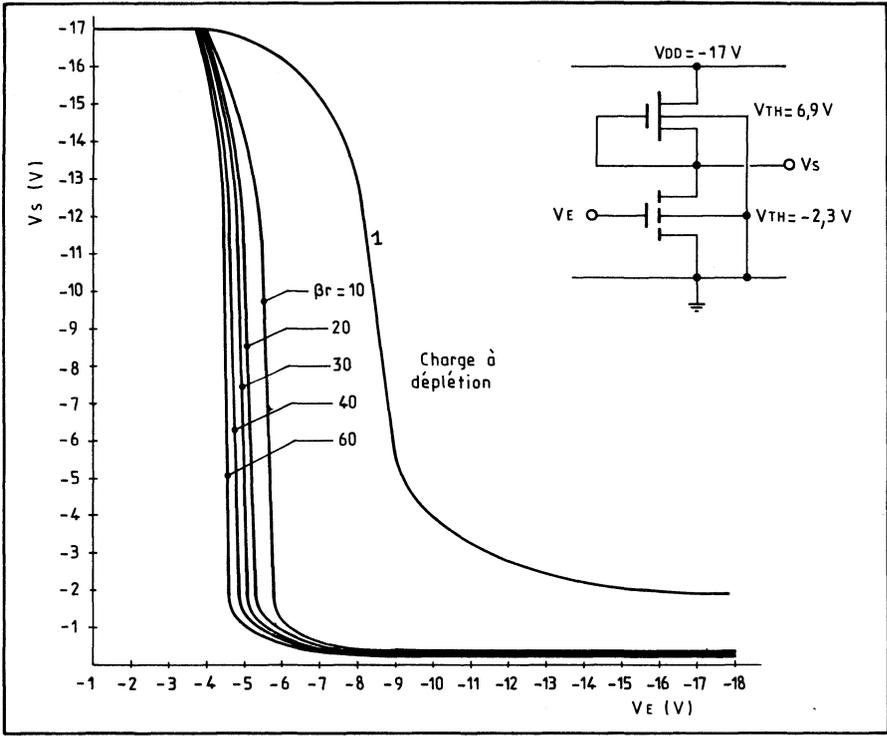


Fig. 13. — Fonctions de transfert types pour divers rapports W/L charge/inverseur.

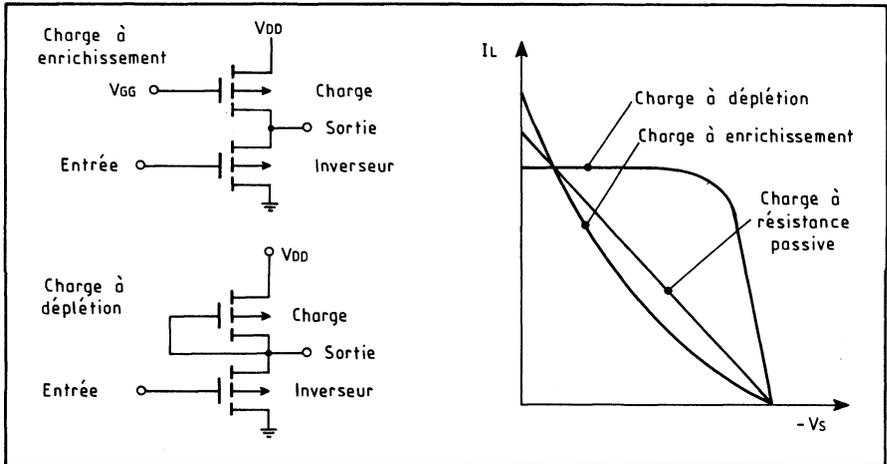


Fig. 14. — Fonctions de transfert comparées pour divers montages de base.

La charge étant toujours conductrice, même lorsque la sortie est au niveau bas, le circuit répond plus rapidement lors des commutations. La droite de charge est dessinée figure 14, où elle est comparée à celles des étages inverseurs à résistance passive et MOS non saturé en charge.

5. L'INVERSEUR A MOS COMPLÉMENTAIRES

L'étage inverseur de base à MOS complémentaires, ou CMOS, comprend deux transistors en série : l'un à canal P, l'autre à canal N, tous deux à enrichissement (fig. 15). Le substrat du PMOS est connecté au $+V_{DD}$ et celui du NMOS à la masse; les gates sont attaqués en parallèle, la sortie se faisant au point commun des drains.

On notera qu'ici, V_{DD} est une tension positive par rapport à la masse, pour respecter les conditions normales de fonctionnement des MOS. Partant, on raisonne en *logique positive*.

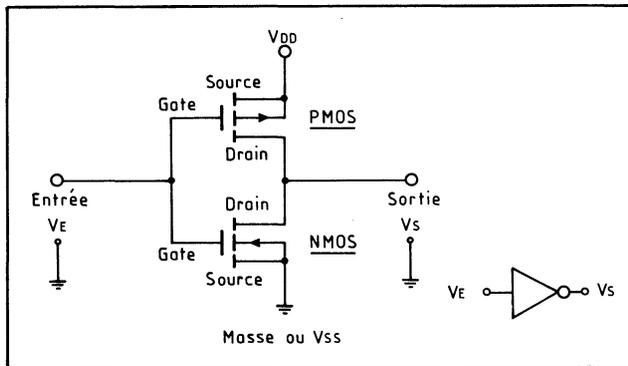


Fig. 15. — L'inverseur à MOS complémentaires.

Le fonctionnement de l'étage est le suivant :

- pour un 0 à l'entrée : le PMOS est conducteur et le NMOS est bloqué. La sortie passe donc à V_{DD} ;
- pour un 1 à l'entrée : c'est l'inverse, car le PMOS se bloque tandis que le NMOS devient conducteur. La sortie est à la masse.

Ce fonctionnement est donc tel qu'on peut assimiler l'étage à un double interrupteur en série (fig. 16). On constate que, lors des états stables, c'est-à-dire en dehors des périodes de commutation, il n'existe aucun chemin conducteur entre V_{DD} et masse, ce qui signifie qu'à un état stable (au repos), l'étage a une consommation nulle.

C'est la raison pour laquelle la famille CMOS bat tous les records en économie de consommation. Celle-ci est réduite, en statique, aux inévitables mais excessivement faibles courants de fuite qui se mesurent en nanoampères; on a :

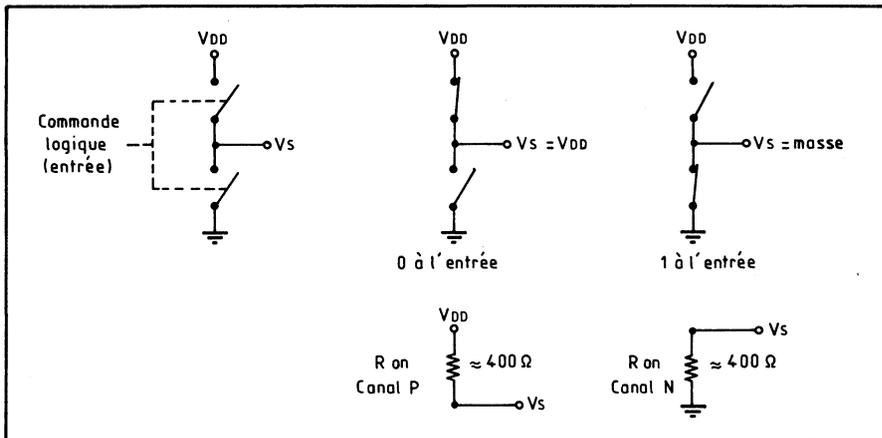


Fig. 16. — Analogie de fonctionnement de la porte CMOS.

$$P = V_{DD} I_F$$

ce qui, pour $V_{DD} = 10 \text{ V}$ et un courant de fuite $I_F = 1 \text{ nA}$, donnerait une consommation de :

$$P = 10 \cdot 10^{-9} = 10 \text{ nanowatts!}$$

La fonction idéale de transfert de la porte inverseuse CMOS est donnée figure 17. La tension de sortie présente une excursion maximale, de la masse à V_{DD} ; la commutation est centrée à mi-courbe; on a indiqué, sur la figure, les valeurs des tensions de seuil V_{TH} pour les MOS, canal N ou P.

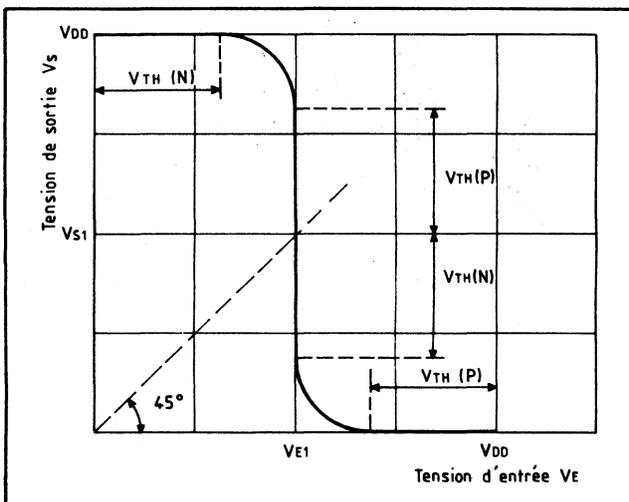


Fig. 17. — Fonction de transfert idéale CMOS.

Fig. 18. — Fonctions de transfert CMOS réelles (porte CD 4001).

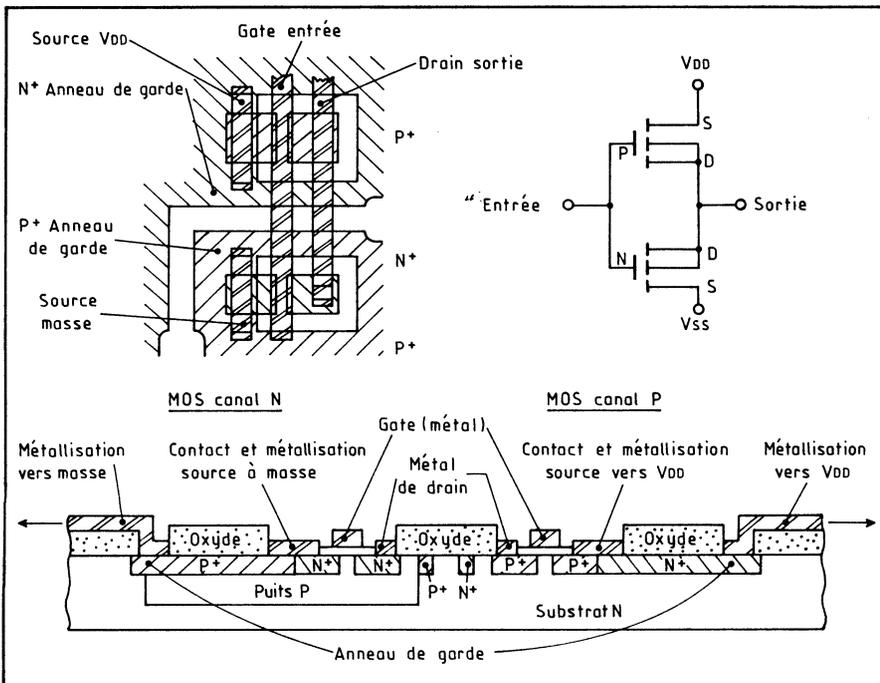
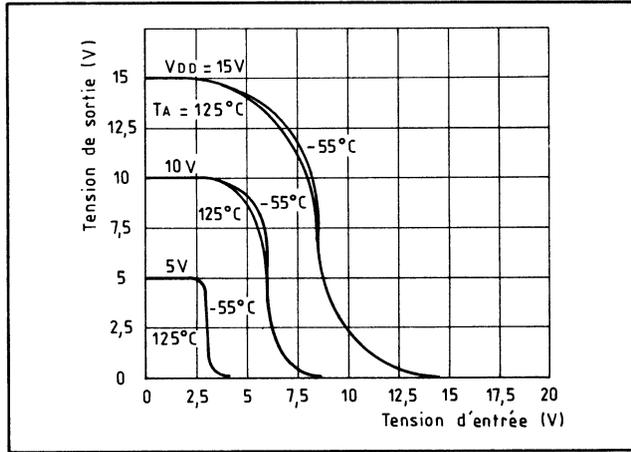


Fig. 19. — Topologie d'un inverseur CMOS.

En réalité, les courbes de transfert ne sont pas aussi parfaites, comme le montre la figure 18; ces courbes ont été relevées sur une porte CD 4001 de RCA avec des tensions V_{DD} de 5, 10 et 15 V. La symétrie des deux MOS n'est pas rigoureuse et souvent, $V_{TH}(P)$ sera supérieur à $V_{TH}(N)$; d'autre part, la tension de sortie n'est égale à V_{DD} où à la masse qu'à 10 mV près environ, chute de tension due au canal conducteur.

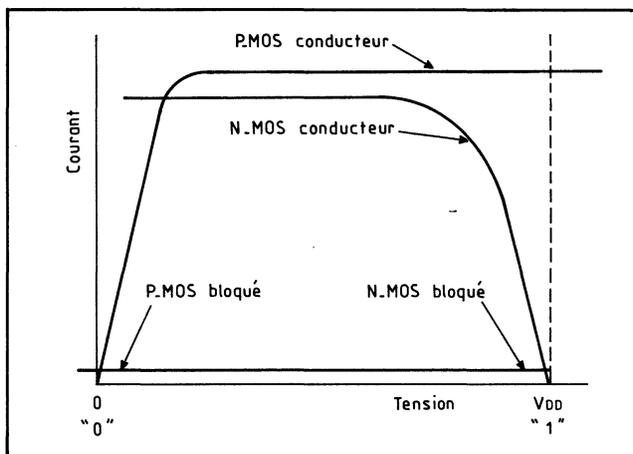


Fig. 20. — Courbes de fonctionnement des PMOS et NMOS.

Le point central de commutation se situe entre 45 et 50 % de V_{DD} , et dépend surtout du rapport des impédances des MOS. Enfin, la température, de -55 à $+125$ °C, n'agit que très peu sur le comportement du système.

Dans la figure 19, on a représenté la topologie d'un inverseur CMOS, et dans la figure 20, l'allure des courants et tensions pour les N et PMOS en commutation.

Fonctionnement en commutation : la consommation

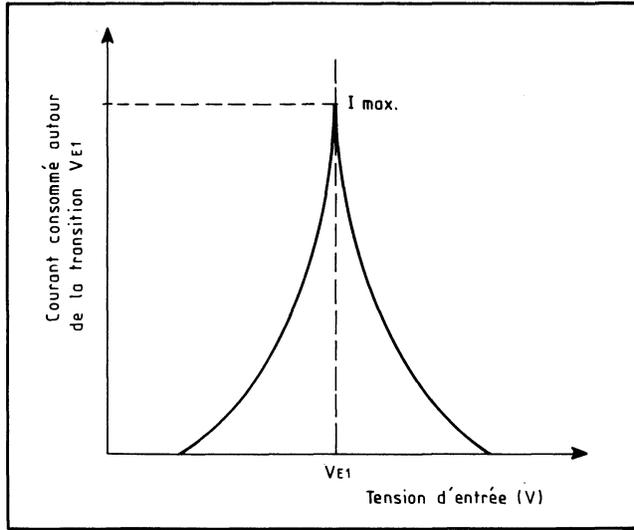
En se référant aux courbes de la fonction de transfert, on voit que pendant la commutation d'un état à un autre, les deux MOS de l'inverseur sont un instant simultanément conducteurs. Par conséquent, et pendant cette transition, une pointe de courant apparaît (fig. 21).

Ce courant crête de commutation, relevé figure 22 sur une porte CD 4011, mène à une consommation en puissance d'autant plus faible que :

1. La tension V_{DD} est plus faible.
2. La commutation est plus rapide.
3. L'impédance des MOS est plus grande.
4. Enfin, que le nombre de commutations par seconde est moindre.

La consommation est ainsi directement proportionnelle à la fréquence F des commutations, au carré de la tension d'alimentation V_{DD} , et à la valeur de la

Fig. 21. — Ce n'est qu'au moment où l'inverseur à CMOS bascule qu'un courant est réellement consommé par l'étage.



capacité C qui charge l'étage. En toute rigueur, la consommation de l'étage, en service réel, est donc la somme de :

- La consommation au repos ;
- La puissance absorbée par les pointes de courant à la commutation ;
- La puissance débitée par l'étage pour charger la capacité de charge sur laquelle il débite.

Or, ces deux premiers termes apparaissent généralement négligeables, eu égard au troisième. Il reste alors :

$$P_{\text{commutation}} = C \cdot V_{DD}^2 \cdot F$$

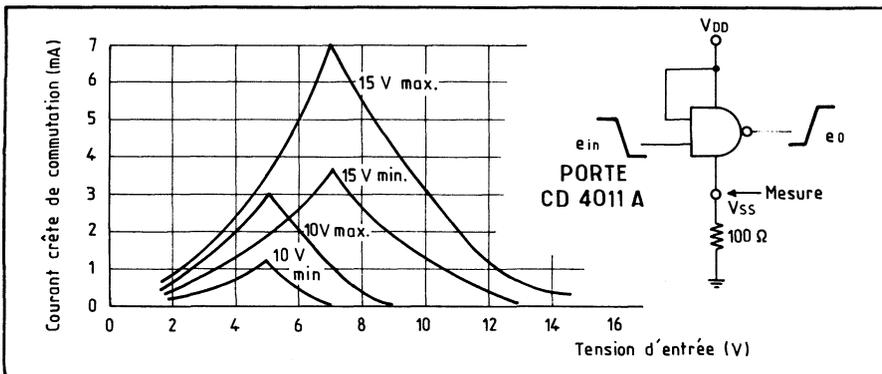


Fig. 22. — Consommation à la commutation d'une porte CMOS type 4011 A.

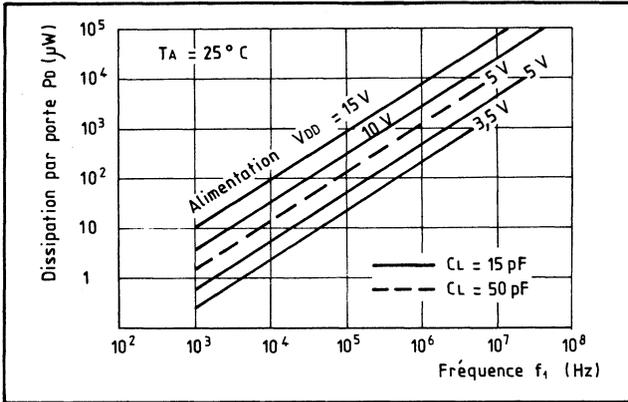


Fig. 23. — Consommation-type en fonction de la fréquence.

Les courbes de la figure 23 montrent comment évolue cette consommation avec la fréquence et pour diverses tensions d'alimentation. A 1 MHz et sous $V_{DD} = 10\text{ V}$, la consommation est de l'ordre de 2 mW, donc très semblable à celle de la famille bipolaire Schottky faible puissance. Mais elle diminue rapidement en dessous de cette fréquence.

La figure 24 donne l'allure des signaux d'entrée (un créneau), de sortie, et du courant V_{DD} à masse qui s'ensuit, avec les principales définitions : *temps de montée*, *temps de chute* (ou de *descente*), *temps de propagation* (t_{pd}) aux niveaux 0 et 1.

Le temps de montée dépend de la charge capacitive (fig. 25), tout comme le temps de descente d'ailleurs (fig. 26). Ces temps sont relativement importants, ce qui contribue à améliorer l'immunité au bruit.

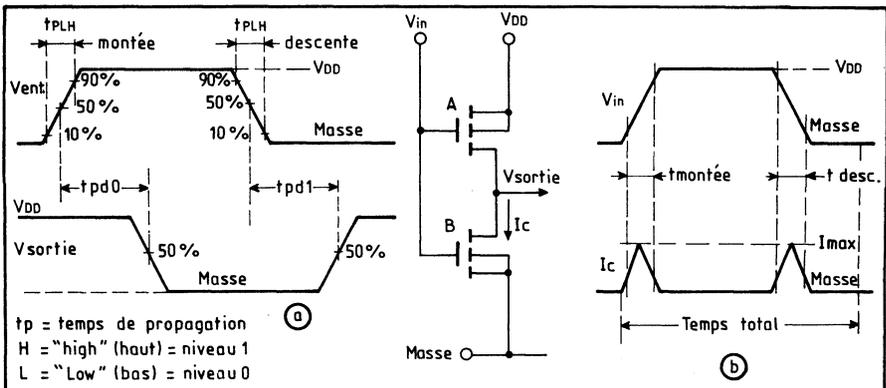


Fig. 24. — Les signaux caractéristiques et l'inverseur CMOS.

Fig. 25. — Temps de montée typique en fonction de la capacité de charge.

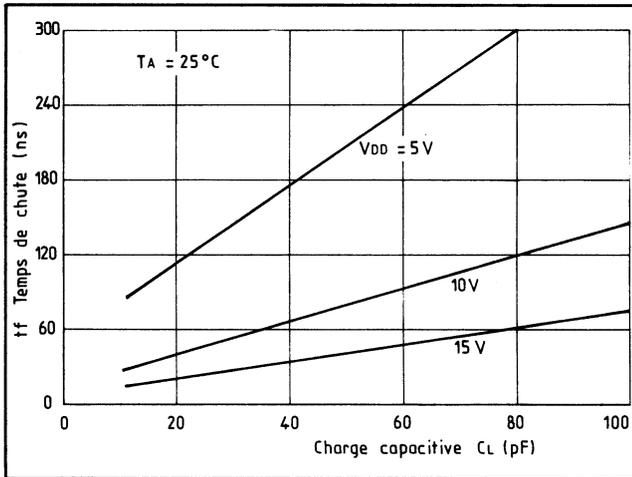
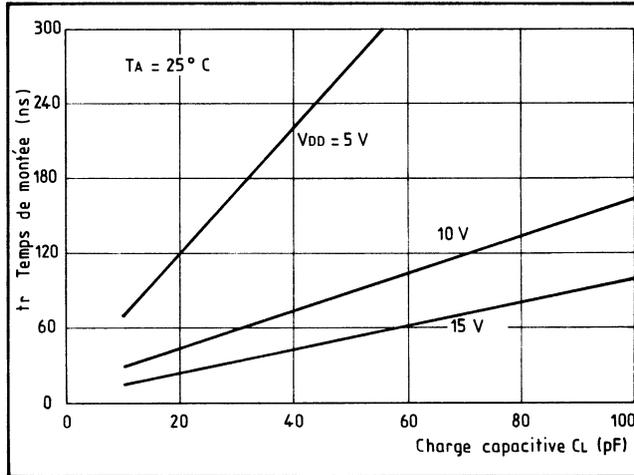


Fig. 26. — Temps de chute typique en fonction de la charge capacitive.

On ne confondra pas les temps de montée (t_r , de *rise time*) et de chute (t_f , de *fall time*) avec le temps de *mise en conduction* ou de *blocage* qui sont plus courts (courbes des fig. 27).

Le temps de propagation diminue lorsque V_{DD} augmente car simultanément, l'impédance du canal conducteur décroît; c'est ce que montrent les courbes de la figure 28, relevées sur une porte CD 4001 et donnant les courants en fonction de la tension pour les MOS à canal P et N.

Le rôle de V_{DD} si l'on veut réduire la consommation

Comme on l'a dit, les pointes de courant de commutation dépendent de la valeur de la tension d'alimentation V_{DD} . Celle-ci détermine également les niveaux

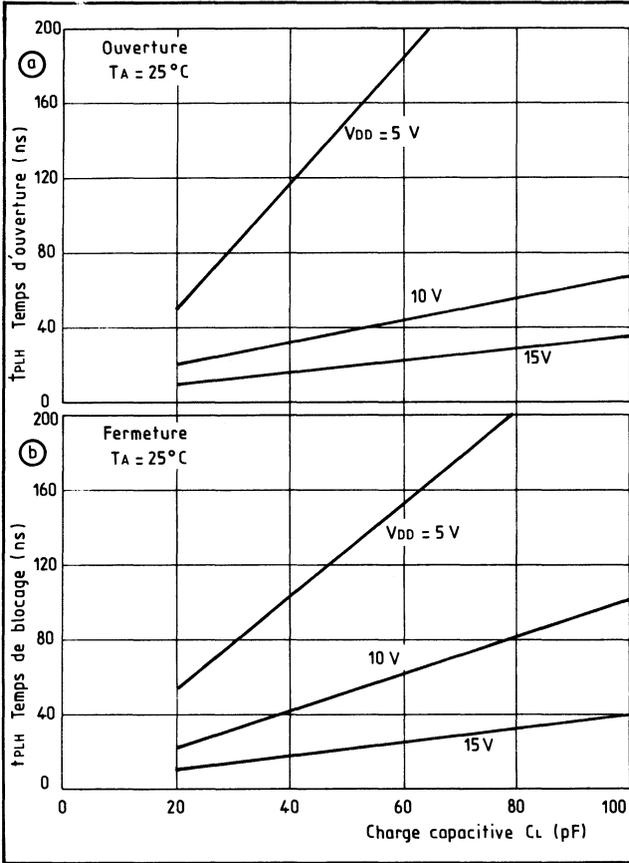


Fig. 27. — Temps d'ouverture (déblocage) et temps de fermeture (blocage) typiques en fonction de la charge.

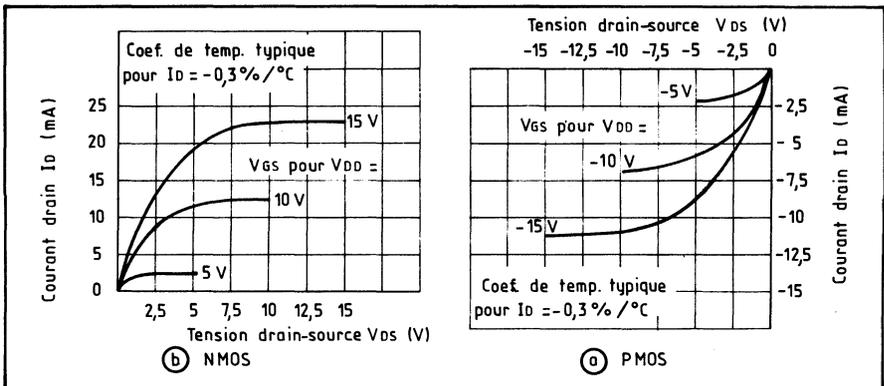
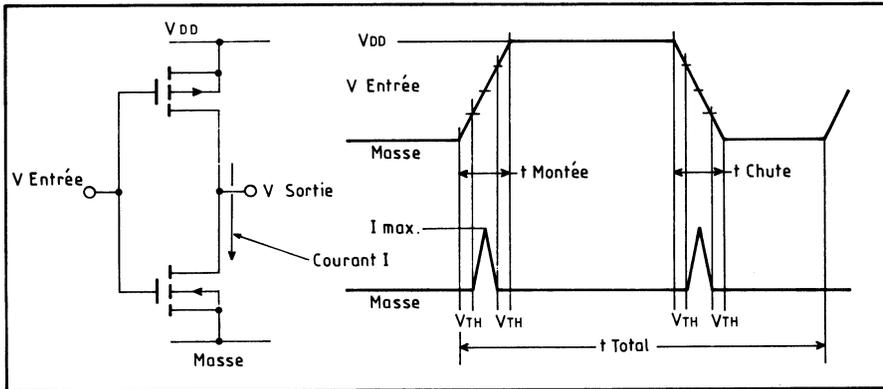
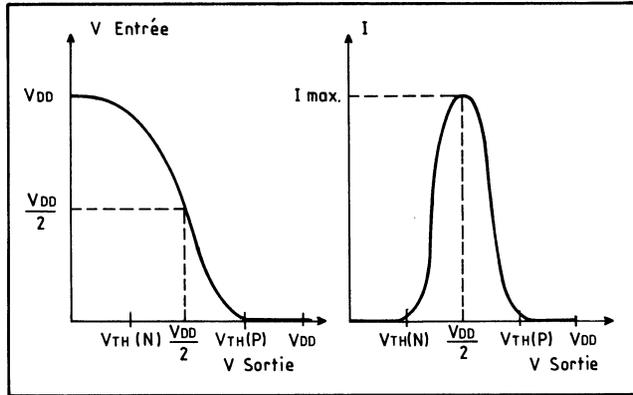


Fig. 28. Caractéristiques typiques des PMOS et NMOS d'une porte complémentaire.

d'attaque des MOS, dans une chaîne d'inverseurs dont le précédent attaque le suivant.

Comme le montre la figure 29, les pointes de courant s'insèrent entre les deux tensions de seuil, du PMOS et du NMOS. Par rapport à ce schéma, en effet, la tension d'attaque s'accroît d'abord le NMOS, puis continue à s'accroître; avant qu'elle atteigne le seuil du PMOS, les deux transistors sont simultanément conducteurs. Elle atteint enfin V_{TH} du PMOS qui se bloque enfin.

Fig. 29. — Caractéristique de transfert et consommation stylisées de l'inverseur CMOS.



Si l'on réduit alors V_{DD} et si l'on suppose que les tensions de seuil V_{TH} des N et PMOS sont identiques, on constate que :

1. Si $V_{DD} > 2 V_{TH}$, la pointe de courant se manifeste à la commutation car les deux MOS sont simultanément conducteurs.
2. Si $V_{DD} = 2 V_{TH}$, un MOS se bloque au moment précis où le second entre en conduction.
3. Si $V_{TH} < V_{DD} < 2 V_{TH}$, un seul MOS peut conduire, mais jamais les deux simultanément.
4. Si $V_{DD} < V_{TH}$, aucun MOS ne peut conduire.

Tableau B. — Immunités au bruit comparées

FAMILLE (porte type)	ALIMENTATION (V)	CONSUMMATION AU REPOS (mW)	TEMPS DE PROPAGATION (ns)		IMMUNITÉ EN TENSION CONTINUE — SIGNAL (V)				IMMUNITÉ TYPIQUE EN ALTER -NATIF		IMPÉDANCE DES LIGNES-SIGNAL (Ω)		IMMUNITÉ TYPIQUE EN ÉNERGIE (nJ) POUR UNE DURÉE D'IMPULSION Dimp (ns)					
					V _{NL}		V _{NH}		ALIMENT. (V)				ÉTATS:					
									+	Mosse			L	H	L		H	
					t _{PHL}	t _{PLH}	Min	Typ	Min	Typ					E _{NL} (nJ)	Dimp (ns)	E _{NL} (nJ)	Dimp (ns)
TTL(MC7400)	5	10	8	12	0,4	1,2	0,4	2,2	3	1	30	140	1,7	20	1	25		
CMOS (MC14011)	5	25 · 10 ⁻⁶	35	100	1,5	2,2	1,5	3,4	2,8	1	1,7k	4,8k	1	155	0,9	280		
	10	50 · 10 ⁻⁶	20	35	3	4,2	3	6	5,7	4,3	670	1,5k	3,7	70	3,1	90		
	15	150 · 10 ⁻⁶	8	15	4,5	6,3	4,5	9	8,5	6,4	460	1k	7,2	50	8,5	75		

**Tableau C. — Quelques configurations
d'inverseurs**

INVERSEUR	CHARGE
Canal P, enrichissement, triode	Résistance passive Canal P, enrichissement, triode Canal P, déplétion, saturation Canal P, enrichissement, penthode Canal N, enrichissement, triode
Canal N, enrichissement, triode	Résistance passive Canal N, enrichissement, triode Canal N, enrichissement, penthode Canal N, déplétion, saturation Canal P, enrichissement, triode
Canal N ou P • Enrichissement, penthode • Déplétion, saturé	Toutes les variantes
MOS complémentaires	

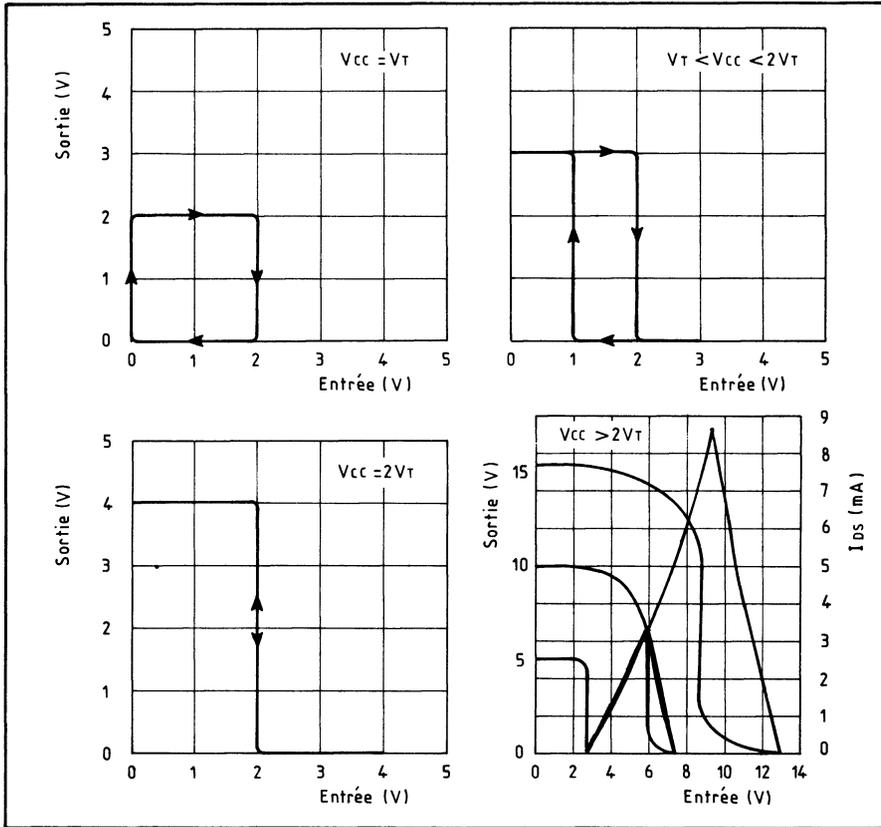


Fig. 30. — Fonctions de transfert selon valeurs de tensions de seuil et d'alimentation pour l'inverseur CMOS.

C'est ce qu'illustrent les courbes de la figure 30 où l'on a fait $V_{TH} = 2\text{ V}$ par souci de simplification.

Si donc on recherche la plus grande économie en consommation, on pourra adopter le cas (3), où V_{DD} est comprise entre les deux tensions de seuil. Tel est le cas en horlogerie où, avec une pile de 1,3 V, on choisit des tensions de seuil telles que $V_{TH}(P) = 0,6\text{ V}$ et $V_{TH}(N) = 0,8\text{ V}$, par exemple.

Immunité au bruit

La fonction de transfert de l'inverseur montre que l'immunité au bruit est excellente; le point central de commutation se situant à environ 45 % de V_{DD} , l'immunité est de 45 % de V_{DD} .

Ainsi, avec $V_{DD} = 10\text{ V}$, l'immunité est de 4,5 V. Si V_{DD} passe à 5 V, tension d'alimentation typique des familles bipolaires, l'immunité est de 2,25 V (contre moins de 1 V pour la TTL, dans ce cas).

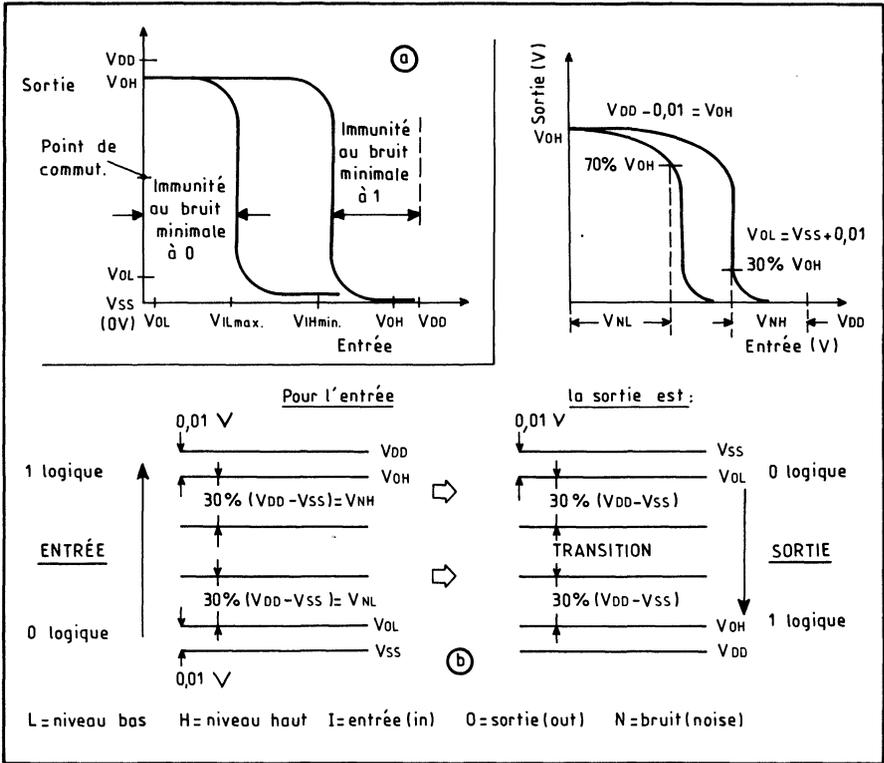


Fig. 31. — Distribution de l'immunité au bruit : L et H précisent les niveaux (L = bas = 0 et H = haut = 1) pour l'entrée (I) et la sortie (O).

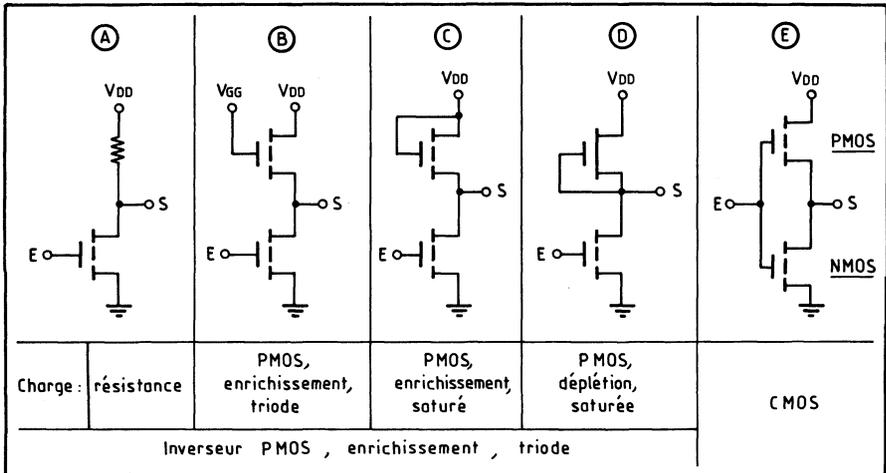


Fig. 32. — Les structures fondamentales de portes inverseuses.

Pratiquement et pour tenir compte de l'inévitable dispersion des caractéristiques, les fabricants garantissent une immunité au bruit égale à 30 % de V_{DD} sous toutes les conditions de température (fig. 31). Le tableau B illustre ce point, en comparant les CMOS à la TTL classique 54/74.

6. SYNTHÈSE DES INVERSEURS STATIQUES

On a examiné ci-dessus plusieurs types d'inverseurs MOS. Or, il en existe bien d'autres, par exemple : NMOS inverseur chargé par un NMOS, tous deux à enrichissement; ou encore NMOS à enrichissement chargé par un NMOS à déplétion...

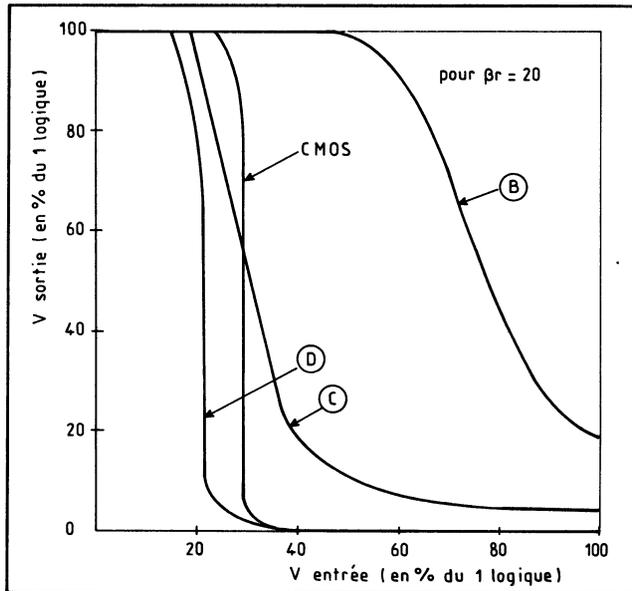


Fig. 33. — Courbes de transfert comparées, relatives aux schémas de la figure précédente.

Les combinaisons peuvent se faire entre PMOS et NMOS, à déplétion ou à enrichissement, fonctionnant en triode ou en saturé, ou même avec résistance passive en charge, et comprennent plusieurs dizaines de variantes d'intérêt inégal. Le tableau C en donne un aperçu. Toutes ces combinaisons sont du type statique, par opposition aux portes dynamiques étudiées plus loin.

Les cinq structures les plus significatives ont été ainsi analysées ci-dessus; elles sont rappelées dans la figure 32 avec leurs courbes de transfert-type (fig. 33).

7. LA CAPACITÉ INTRINSÈQUE (GATE-SUBSTRAT)

L'inverseur fondamental est normalement incorporé dans les montages plus complexes où apparaissent des chaînes d'inverseurs ou de portes. Ainsi, on peut

considérer qu'un inverseur en attaque un second (fig. 34). De ce fait, le rôle de la première porte consiste exclusivement à charger ou décharger la capacité d'entrée gate-substrat, ou *capacité intrinsèque*, de la seconde.

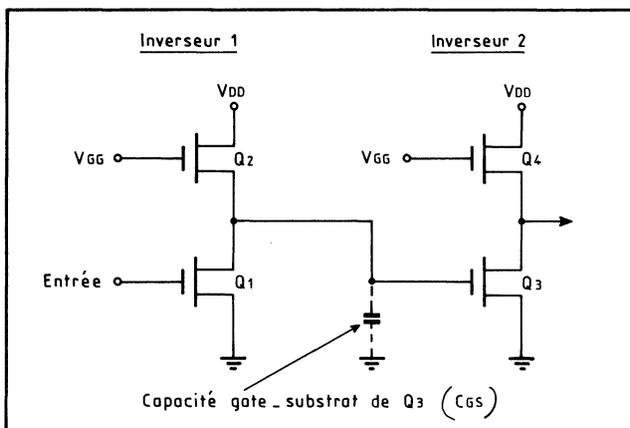


Fig. 34. — Place de la capacité intrinsèque d'entrée C_{GS} .

Le transistor Q_1 n'intervient qu'en fonction des ordres appliqués à l'entrée; par contre, Q_2 est en service en permanence. On peut alors totalement séparer la capacité C_{GS} de Q_2 en intercalant un MOS supplémentaire entre les deux inverseurs (fig. 35). Pratiquement, lorsque ce MOS, noté Q_5 , est bloqué, la capacité C_{GS} est complètement isolée.

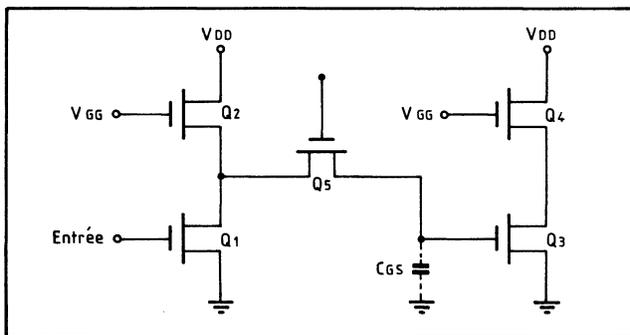


Fig. 35. — Comment l'on sépare et isole les deux inverseurs.

De ce fait, elle conservera l'état de charge qu'on lui aura fourni. En effet, l'impédance d'entrée de Q_3 est énorme et n'intervient pas; seuls, les courants de fuite via Q_5 viendront perturber cette mémorisation. Or, ils sont très faibles et pratiquement, on considère que le stockage est maintenu sans altération pendant plusieurs millisecondes (2 à 5). Tel est le principe du point mémoire « dynamique », à MOS.

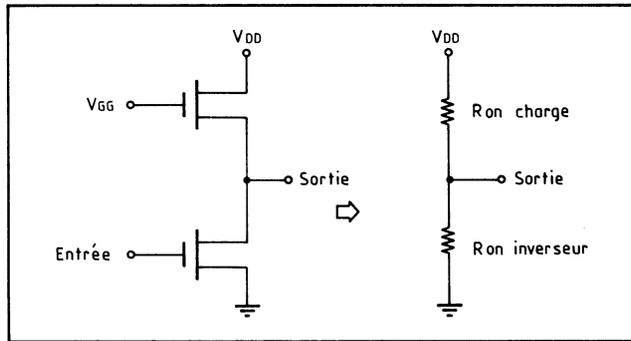
8. LA PORTE DE TRANSFERT

Le transistor Q_5 du montage précédent, lui, en raison de sa fonction, est appelé *porte de transfert*. Il confère au montage des propriétés nouvelles exploitant la propriété des MOS d'être des éléments bidirectionnels, ainsi qu'on le verra ultérieurement de façon plus détaillée.

9. LES INVERSEURS DYNAMIQUES

Les portes analysées précédemment offrent donc toutes un trait commun : elles fonctionnent en régime continu. Or, on peut imaginer un fonctionnement en impulsions qui mette à profit la grande impédance des MOS bloqués et, par voie de conséquence, l'existence d'une capacité intrinsèque gate-substrat. Un tel fonctionnement est dit « dynamique », par opposition au fonctionnement « statique » qui qualifie les montages déjà étudiés.

Fig. 36. — Principe du montage proportionnel.



On va donc examiner ici le principe des inverseurs dynamiques; pourtant, on notera qu'ils ne sont pas utilisés tels, isolément, mais qu'ils participent à des montages plus complexes (registres à décalage, mémoires RAM dynamiques) où leurs qualités sont réellement exploitées. Parmi celles-ci on relèvera, selon les circuits :

- une plus grande simplicité des schémas dans certains cas (mémoires RAM...);
- une plus grande économie en consommation;
- une densité d'intégration accrue (elle résulte des deux précédents avantages).

En fait, ces portes dynamiques se partagent en *portes proportionnelles* et en *portes non proportionnelles*.

Inverseur proportionnel

L'inverseur de base à MOS fournit une tension de sortie, au niveau *bas*, proportionnelle à la résistance des canaux conducteurs (R_{on}) de l'inverseur proprement dit et de la charge. C'est ce que montre la figure 36. Ce même raisonnement vaut d'ailleurs lorsque la sortie est au niveau haut. De là provient le qualificatif *d'inverseur proportionnel*.

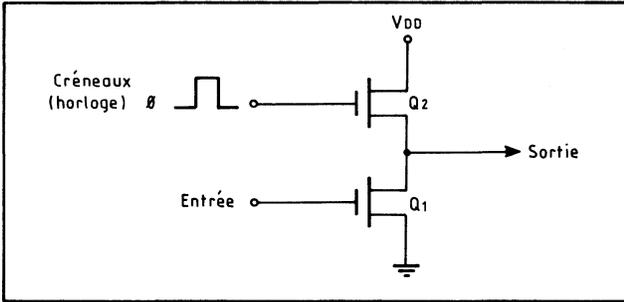


Fig. 37. — Porte synchrone.

Lorsque deux portes se suivent, on constate que la première a pour rôle exclusif de charger et décharger la capacité de structure de la seconde. Pour isoler complètement celle-ci, on peut intercaler une porte de transfert, comme on l'a vu, à la sortie de la première porte. On pourrait également imaginer d'attaquer en impulsions le MOS de charge, afin de ne le mettre en service qu'en des périodes bien définies (fig. 37); on réalise ainsi une porte *synchrone*.

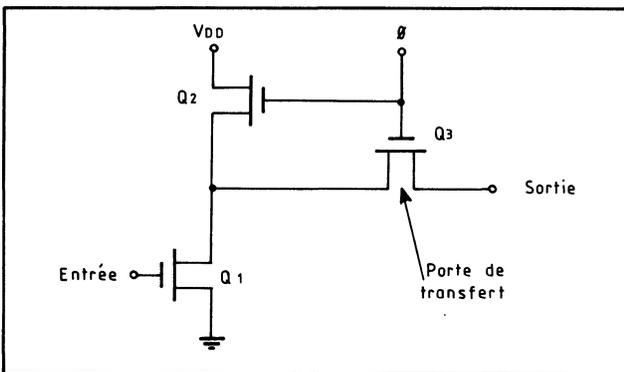


Fig. 38. — Porte dynamique proportionnelle.

Cependant, pour obtenir le fonctionnement le plus satisfaisant, on jumelle souvent ces deux solutions (fig. 38). Dans la structure résultante, les créneaux d'horloge sont appliqués simultanément à Q_2 et Q_3 qu'ils rendent conducteurs; en leur absence, ces transistors sont bloqués et le courant de fuite est réduit au minimum. Une telle porte, à une unique phase d'horloge, servira à réaliser des registres à décalage biphasés et est dite *dynamique proportionnelle*.

Inverseur non proportionnel à précharge

Dans le circuit précédent, un chemin continu entre V_{DD} et masse se manifeste lorsque Q_1 est débloqué en présence du signal d'horloge; il s'agissait donc d'un circuit proportionnel.

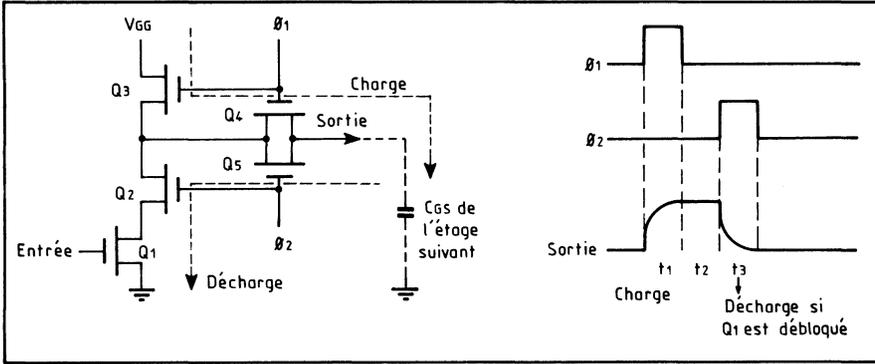


Fig. 39. — Inverseur dynamique non proportionnel à précharge.

Pour couper ce chemin continu, il suffit d'intercaler un transistor supplémentaire dans la chaîne, comme le montre la figure 39, et de recourir cette fois à deux phases d'horloge, Φ_1 et Φ_2 . On est alors contraint de prévoir également deux MOS pour la porte de transfert, afin de séparer les commandes d'horloge.

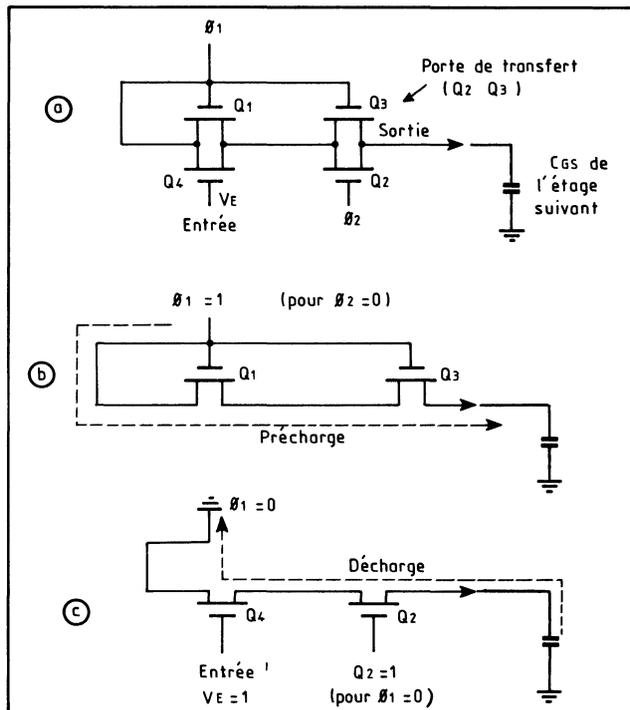


Fig. 40. — Inverseur non proportionnel alimenté par l'horloge (a); détail du fonctionnement selon que l'une ou l'autre phase d'horloge est en service (b et c).

Il est évident que la phase Φ_1 sert alors à charger la capacité C ; celle-ci étant « préchargée » la phase Φ_2 permettra de la décharger si Q_1 est conducteur, sinon, elle restera dans l'état.

On dit d'un tel circuit qu'il est à *précharge*, car cette opération, équivalent à un armement, précède toutes les autres. D'autre part, la chaîne résistive de division a ici disparu et la tension de sortie n'est plus proportionnelle au rapport des résistances R_{on} : ces résistances apparaissent en série (selon les transistors activés); ainsi dit-on qu'il s'agit d'un *inverseur non proportionnel*.

Encore faut-il que les créneaux Φ_1 et Φ_2 fournis par des « horloges », se succèdent dans le temps sans chevauchement, faute de quoi l'on se trouverait ramené au cas précédent.

Un tel circuit non proportionnel se manifeste par une vitesse accrue, par rapport au montage proportionnel. Avec des MOS de même type, on atteindra 2 à 5 MHz, par exemple, là où l'inverseur statique, proportionnel, plafonne vers le mégahertz.

Inverseur non proportionnel à précharge alimenté par l'horloge

La structure de l'inverseur non proportionnel suivant est un peu plus déroutante (fig. 40 a). On retrouve le transistor inverseur Q_3 ; une porte double de transfert Q_2 et Q_3 ; ainsi qu'un MOS pour l'horloge Φ_1 et la précharge. On suppose que les deux horloges, Φ_1 et Φ_2 , peuvent prendre les valeurs 1 et 0, soit V_{GG} et masse. Les cas suivants sont possibles :

1. $\Phi_1 = 1$ alors que $\Phi_2 = 0$ et $V_E = 0$: les transistors Q_1 et Q_2 conduisent et chargent C_{GS} (en b sur la figure).

2. $\Phi_1 = 0$ alors que $\Phi_2 = 1$ et $V_E = 1$: les transistors Q_1 et Q_3 sont bloqués, mais Q_2 et Q_4 conduisent, donc C_{GS} se décharge à la masse, c'est-à-dire à Φ_1 (en c).

3. $\Phi_1 = 0$ et $\Phi_2 = 1$, comme précédemment, mais $V_E = 0$: il n'existe aucun chemin de décharge.

4. Dernier cas intéressant, Φ_1 et Φ_2 sont simultanément à 1, Q_4 étant ou non bloqué : la capacité C_{GS} se charge comme en 1.

Il n'y a donc plus ici, de contraintes pour les phases d'horloge Φ_1 et Φ_2 qui peuvent se chevaucher. Le tableau résume toutes les situations :

Φ_1	Φ_2	V_E	Conducteurs	Bloqués	C_{GS}
0	0	0	Aucun	Tous	} Sans changement
0	0	1	Q_4	Q_1, Q_2, Q_3	
1	0	0	Q_1, Q_3	Q_2, Q_4	} Se charge
1	0	1	Q_1, Q_3, Q_4	Q_2	
0	1	0	Q_2	Q_1, Q_3, Q_4	Sans changement
0	1	1	Q_2, Q_4	Q_1, Q_3	Se décharge
1	1	0	Q_1, Q_2, Q_3	Q_4	} Se charge
1	1	1	Tous	Aucun	

Ainsi, ce sont ces horloges qui fournissent l'alimentation, masse et V_{GG} , ou plus exactement la phase Φ_1 ; en conséquence, celle-ci doit être capable de fournir la puissance nécessaire. Par contre, on économise les lignes de masse et de V_{GG} classiques. Enfin, et par rapport à l'inverseur non proportionnel précédent, on gagne un MOS par inverseur. Cet inverseur biphasé, non proportionnel, est dit *à alimentation par les horloges*.

CHAPITRE X

LES PORTES LOGIQUES

1. LES PORTES ET LEURS CARACTÉRISTIQUES

On a examiné dans le chapitre précédent le fonctionnement de l'inverseur, et celui de la porte de transfert. Les autres portes sont : le AND (*ET logique*) et le OR (*OU logique*), ces mêmes fonctions avec inverseur, c'est-à-dire les NAND (*non-ET*) et NOR (*non-OU*, appelé *NI*), ainsi que des structures qui en dérivent et dont les principales sont le OU exclusif noté EXOR, ainsi que le ET exclusif, qui détecte une coïncidence et est appelé COIN.

Le tableau A en résume les états, à partir de portes à 2 entrées et ce, en logique négative (ou positive, en particulier pour les CMOS). Cependant, on se rappellera que si l'on inverse les logiques et si l'on passe de la négative à la positive, ou vice-versa, les fonctions changent et le NOR, par exemple, devient un NAND (ce qui apparaît nettement à la lecture du tableau si l'on remplace les 1 par des 0 et les 0 par des 1).

**Tableau A. — Tables d'états
des principales portes (pour 2 entrées A et B)**

ENTRÉES		SORTIE					
A	B	AND	OR	NAND	NOR	X OR	COIN
0	0	0	0	1	1	0	1
0	1	0	1	1	0	1	0
1	0	0	1	1	0	1	0
1	1	1	1	0	0	0	1
SORTIE =		AB	$A+B$	\overline{AB}	$\overline{A+B}$	$A\oplus B$	$\overline{A\oplus B}$

Le temps de propagation du signal à travers une porte est défini dans la figure 1. On suppose qu'on applique un créneau à l'entrée d'un inverseur, ou à l'une des entrées d'un NOR, par exemple, puisque cette structure est fondamentale avec les MOS; à la sortie apparaît un créneau, les définitions étant :

t_{PHL} temps de propagation de l'entrée *haute* (« high » = H) à la sortie *basse* (low = L).

t_{PLH} temps de propagation de l'entrée, retournant au niveau *bas* à la sortie revenant au niveau *haut*.

t_r temps de montée (de *rise time*), de 10 à 90 %.

t_f temps de chute (de *fall time*), de 90 à 10 %.

t_{THL} temps de transition à la sortie, *haut vers bas*.

t_{TLH} temps de transition à la sortie, *bas vers haut*.

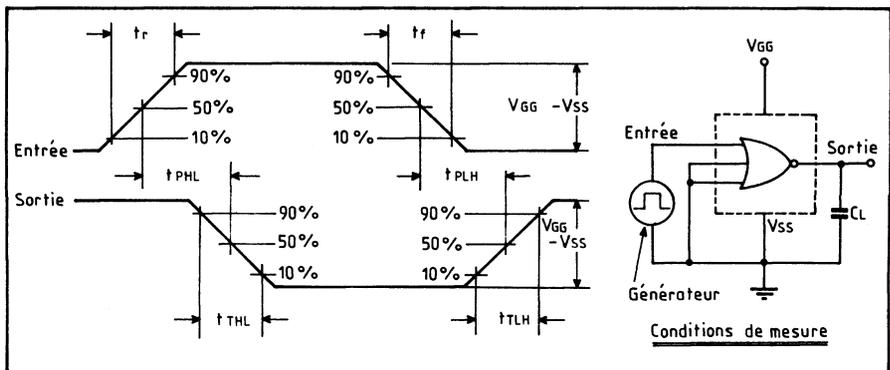


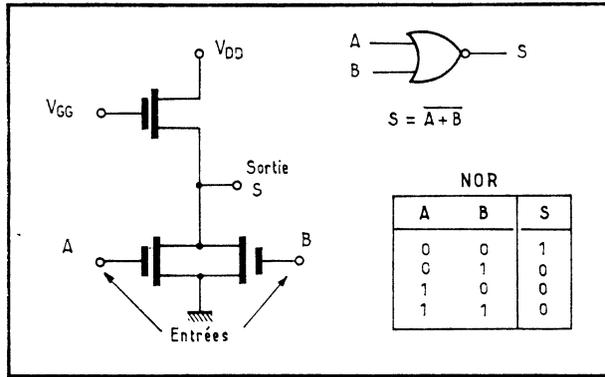
Fig. 1. — Les signaux caractéristiques et les définitions des principaux coefficients.

2. PORTES STATIQUES

Portes de base

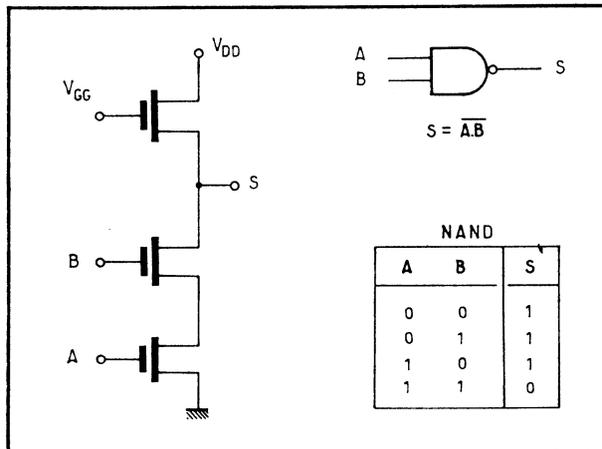
La structure de la porte NOR est extrêmement simple (fig. 2). Le MOS actif de l'inverseur est multiplié, en parallèle, en autant d'autres transistors qu'il y a d'entrées. Il est évident que la sortie ne sera à 1 que lorsque tous ces MOS en parallèle seront bloqués.

Fig. 2. — La porte NOR fondamentale, ici à deux entrées.



Dans le cas du NAND, la structure est également très simple puisqu'on se borne à connecter des MOS en série : autant de MOS qu'il y a d'entrées, plus la charge bien évidemment (fig. 3).

Fig. 3. — Le NAND fondamental, à deux entrées.



La porte AND à deux entrées demande deux MOS de plus que le NAND (fig. 4). En fait, il s'agit tout bonnement d'un NAND suivi d'un inverseur.

De la même façon, le OR est réalisé en ajoutant un inverseur à un NOR fondamental (fig. 5).

Porte OU exclusif (EXOR)

Le montage du OU exclusif (EXOR, ou encore simplement XOR) est un peu plus complexe (fig. 6), bien qu'encore très simple par rapport à son homologue bipolaire. La porte NOR constituée par Q_1 et Q_2 attaque le gate de Q_7 , monté en

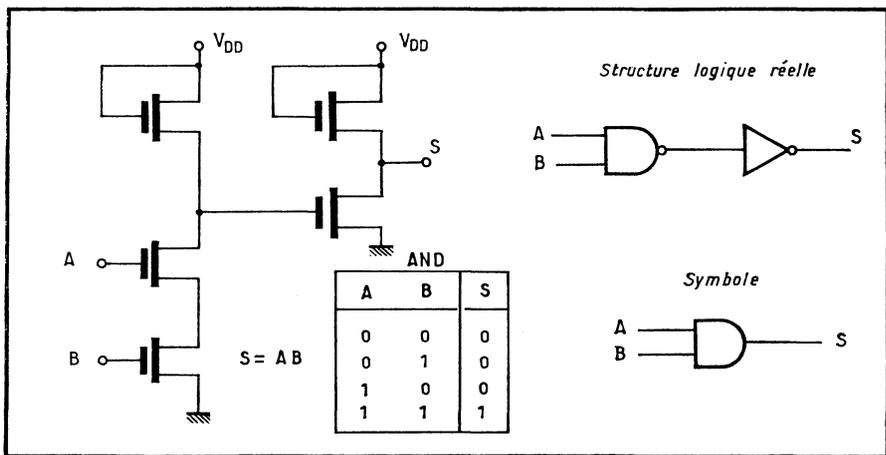


Fig. 4. — Le AND se compose, en fait, d'un NAND suivi d'un inverseur.

NOR également avec la chaîne série de Q_4 et Q_5 , qui peuvent alors être considérés, isolément, comme un AND. Les MOS notés Q_3 et Q_6 sont les charges. Ainsi :

1. Pour $A = 0$ et $B = 0$, le NAND de Q_4 et Q_5 porte la sortie à 0.
2. Pour $A = 1$ et $B = 1$, c'est le NOR de Q_1 et Q_2 qui intervient en portant le point G à 1, ce qui a pour effet de débloquent Q_7 ; par conséquent, la sortie passe à 0.
3. Par contre, pour $A = 1$ et $B = 0$, ou pour l'inverse, $A = 0$ et $B = 1$, le NOR de Q_1 et Q_2 est débloquent et en conséquence, Q_7 est bloqué, alors que le NAND de Q_4 et Q_5 reste fermé et commande le niveau de sortie qui passe à 1.

Le schéma logique correspondant est donné dans la même figure, avec l'équation de sortie (qui peut être transposée sous d'autres formes plus connues du OU exclusif). Un autre montage équivalent est donné en *d*.

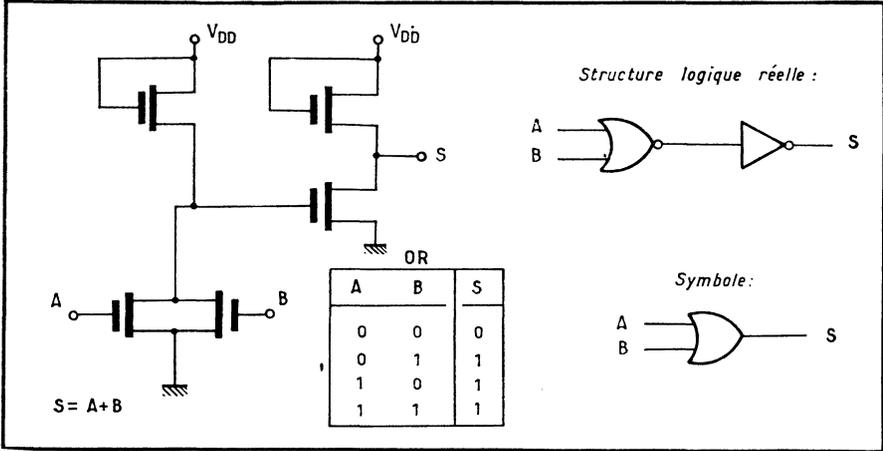


Fig. 5. — Le OR se compose, lui, d'un NOR suivi d'un inverseur.

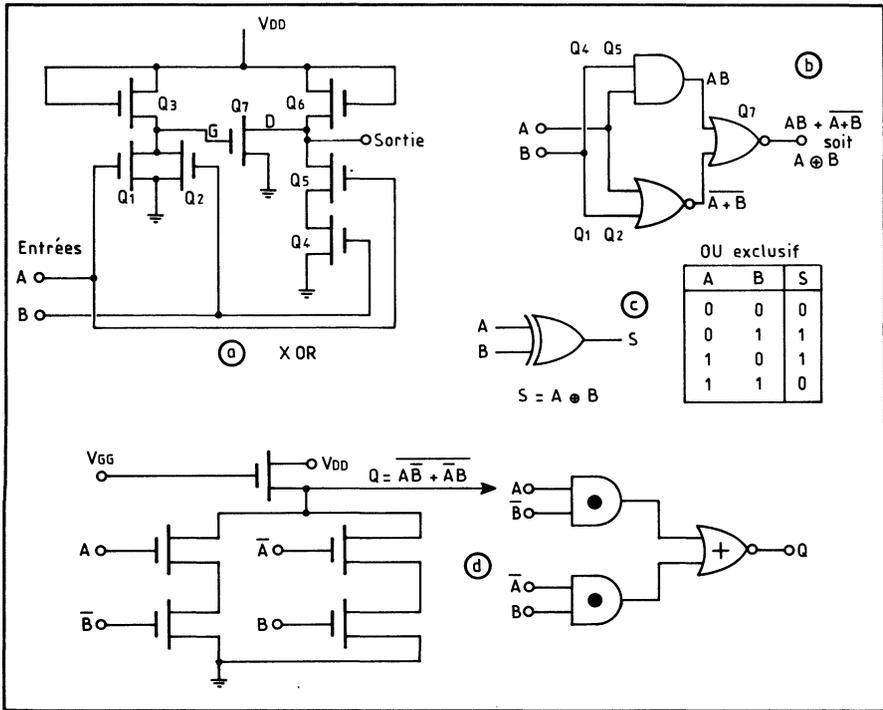


Fig. 6. — Le OU exclusif; a, b et c : caractéristiques; en d : variante électrique et logique.

Portes diverses

Comme le EXOR l'a montré, les combinaisons les plus variées sont possibles. Ainsi, la figure 7 reprend une partie du EXOR sous sa forme ET/OU/NON, et la figure 8 montre le OR/NAND et le AND/NOR.

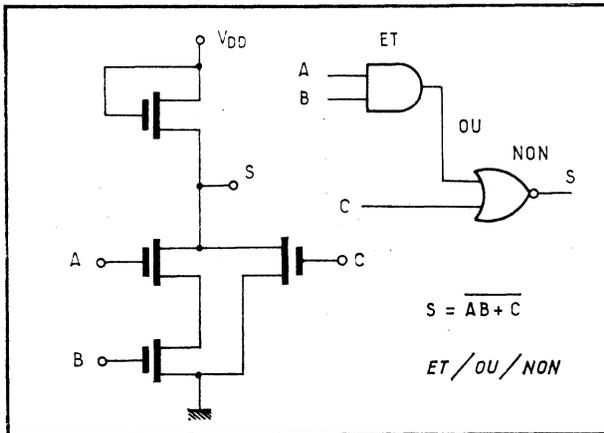


Fig. 7. — Porte ET/OU/NON

Là où de la puissance est exigée, pour assurer la commande de lignes ou de charges capacitives, par exemple, on utilise une porte de sortie de puissance du type push-pull (fig. 9).

Grâce à l'inverseur, l'attaque de la porte de sortie est assurée en symétrique et lorsque Q_3 est conducteur, Q_4 est bel et bien bloqué. Souvent, la porte de puissance est appelée « buffer », ce qui désigne aussi bien d'ailleurs une porte-tampon.

3. LE OU CABLÉ ET LA SORTIE 3 ÉTATS

On dit d'un circuit — inverseur, porte, bascule ou autre — qu'il est à sortie 3 états lorsque la sortie peut prendre trois états :

1. Le niveau *bas* normal (0);
2. Le niveau *haut* normal (1);
3. Un niveau indéterminé, à haute impédance, correspondant à une sortie déconnectée du circuit (en l'air).

L'avantage de ce 3^e état, c'est qu'il permet de connecter en parallèle les sorties de plusieurs dizaines de circuits dont un seul est mis en service à la fois. C'est là une nécessité lorsqu'on recourt à des *bus* de liaison (fig. 10 a).

En effet, on voit alors que si un seul des MOS inverseurs, T_1 , T_2 ou T_3 , est conducteur, il porte la sortie à 0 mais doit accepter la totalité du courant traversant

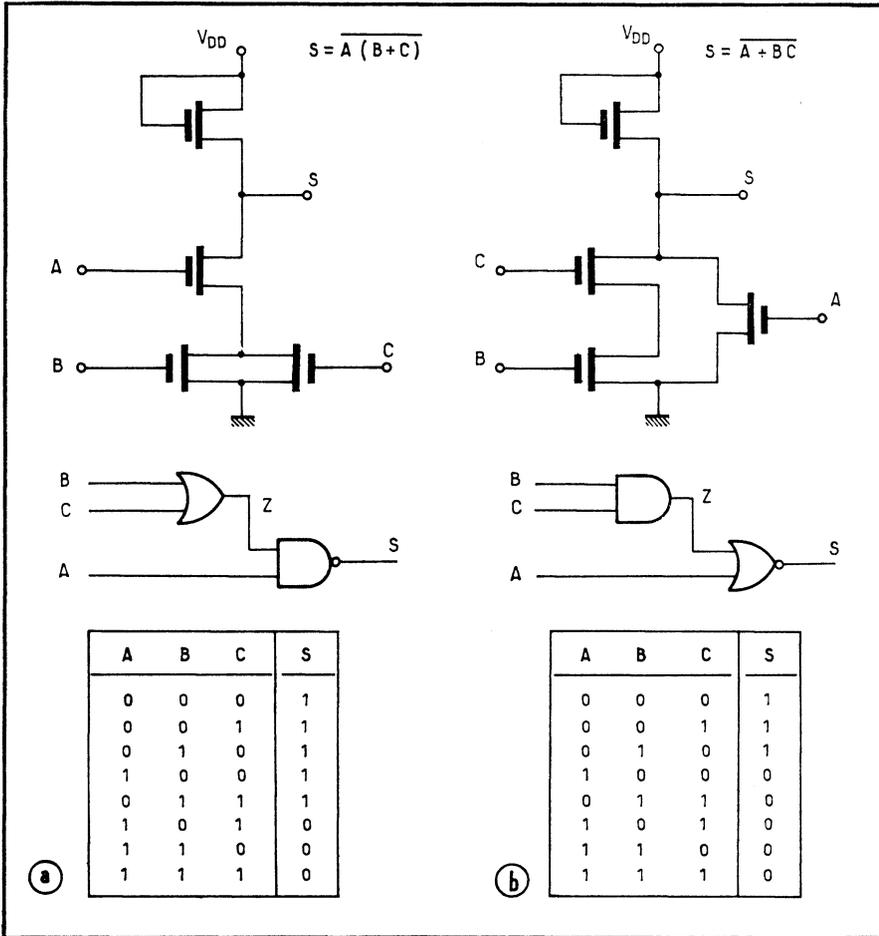
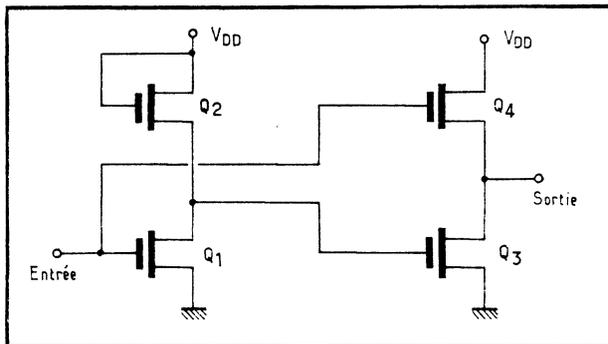


Fig. 8. — Deux portes plus complexes.

Fig. 9. — On utilise fréquemment, en guise de porte de sortie, un montage push-pull constitué par les transistors Q3 et Q4 de cette figure. Cette porte de puissance est, en effet, attaquée en inversion de phase par l'inverseur Q1 et Q2.



les charges en parallèle. Une solution consisterait à utiliser des portes sans charge, et à leur adjoindre une charge commune externe (en *b*).

Cette méthode reste pourtant impraticable avec les MOS complémentaires, pour la même raison qu'avec la TTL bipolaire puisque les sorties de ces deux familles fonctionnent comme de doubles interrupteurs en série. De ce fait, la connexion des sorties à un bus unique est interdite, sous peine de court-circuit lorsque les états ne sont pas les mêmes (en *c*); on voit, en effet, que si T_2 et T_3 sont simultanément conducteurs (ou T_1 et T_4), dans le cas de deux portes, le montage est équivalent à un court-circuit entre V_{DD} et masse.

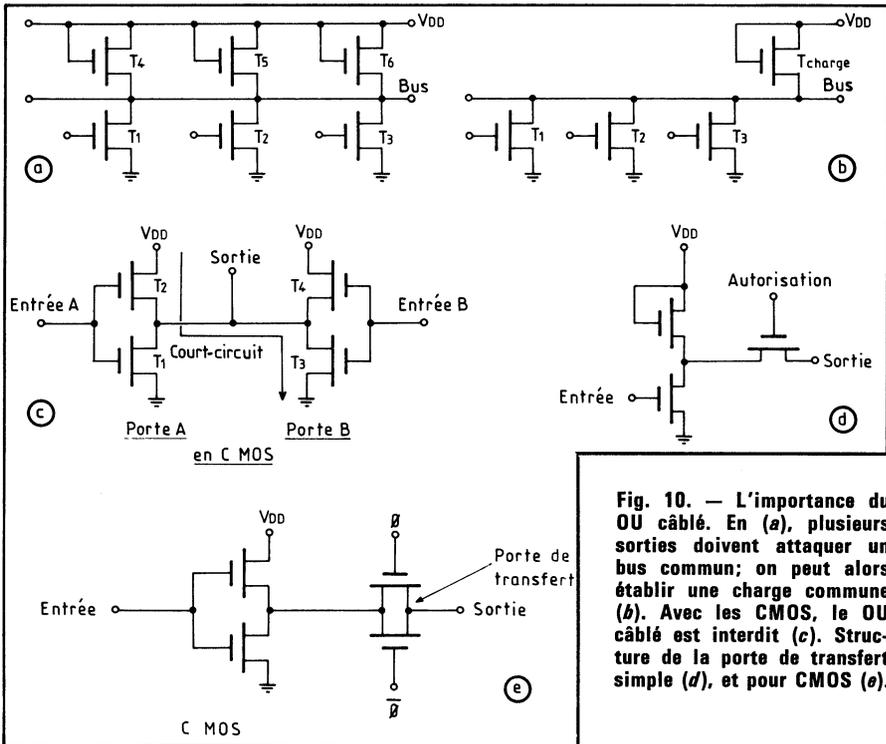


Fig. 10. — L'importance du OU câblé. En (a), plusieurs sorties doivent attaquer un bus commun; on peut alors établir une charge commune (b). Avec les CMOS, le OU câblé est interdit (c). Structure de la porte de transfert simple (d), et pour CMOS (e).

Rappelons qu'une sortie commune à plusieurs circuits exécute la fonction OU, qu'on appelle dans ce cas *OU câblé* (ou *OU connecté*), et elle s'avère très importante dans bien des applications (informatique, mémoires...).

Pour la réaliser, on recourt alors à la formule trois états proposée à l'origine par *National Semiconductor* pour la TTL, mais bien plus aisée à appliquer aux MOS puisqu'il suffit de connecter en série avec chaque sortie une porte de transfert, MOS unique, ou double pour les CMOS (en *d* et *e*). Le signal d'autorisation appliqué à cette porte connecte le circuit à la sortie. En son absence, la sortie reste « en l'air ».

On aboutirait au même résultat en ajoutant deux MOS supplémentaires en série avec ceux de l'inverseur, et commandés comme il se doit (on en verra un exemple ultérieurement).

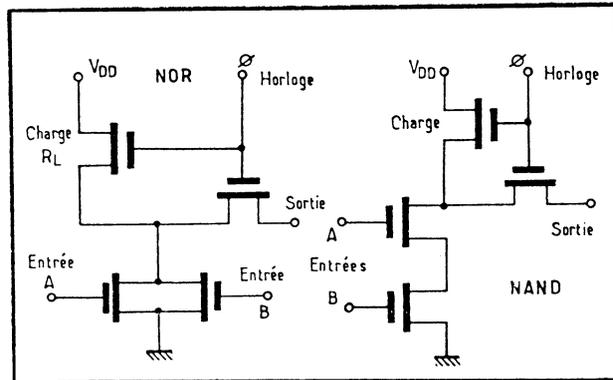
4. PORTES DYNAMIQUES

Les portes dynamiques dérivent simplement des portes statiques, transposées grâce aux montages d'inverseurs dynamiques étudiés antérieurement. Les avantages des portes dynamiques sont, rappelons-le :

- une réduction de la consommation puisqu'on ne travaille plus en régime continu;
- par conséquent, les dimensions des MOS élémentaires peuvent être réduites, puisqu'ils dissipent moins. Il en résulte un accroissement de la densité d'intégration ;
- enfin, les dimensions réduites mènent à leur tour à une diminution des capacités d'entrée. La vitesse s'accroît considérablement ; elle peut être multipliée par un coefficient variant entre 2 et 5. Par exemple, un MOS de petite structure, à R_{on} de 20 k Ω , charge une capacité de 0,5 pF en 10 ns contre 0,1 ms pour 5 pF ;
- ajoutons qu'une réduction des capacités agit à son tour sur la puissance de commande nécessaire, ce qui renvoie en boucle au premier point.

En contrepartie, on s'impose l'utilisation d'horloges et de lignes de synchronisation dont le coût ne semble aisément amorti qu'avec des systèmes plus importants.

Fig. 11. — Portes dynamiques en logique proportionnelle.



Deux portes proportionnelles synchrones sont représentées dans la figure 11 : un NOR et un NAND, sur lesquels il n'est pas utile de faire des commentaires.

Les portes non proportionnelles, à précharge, sont représentées figure 12 dans la version à lignes d'alimentation réelles, et 13 dans la version à alimentation par les horloges, ou 14, où les phases d'horloge sont en série.

5. PORTES CMOS

Le schéma de l'inverseur fondamental en MOS complémentaires, ou CMOS, est rappelé figure 15, avec sa fonction de transfert typique. Quand une tension nulle (zéro logique) est appliquée à l'entrée, le V_{GS} du MOS à canal P est égal à la tension d'alimentation V_{DD} ; de ce fait, ce MOS est conducteur. A l'inverse, le V_{GS} du MOS à canal N est nul, et celui-ci est bloqué. La sortie est donc à 1.

Si un 1 est appliqué à l'entrée, un raisonnement semblable montre que la sortie passe à 0. Puisqu'un transistor est toujours bloqué (à un état stable, hors la commutation), la consommation est des plus réduites. Typiquement, elle est de 1 nA avec $V_{DD} = +10$ V. Elle s'accroît en commutation comme on l'a signalé plus haut de :

$$C_0 (V_{DD})^2 F$$

où C_0 est la capacité de la charge et F la fréquence des commutations.

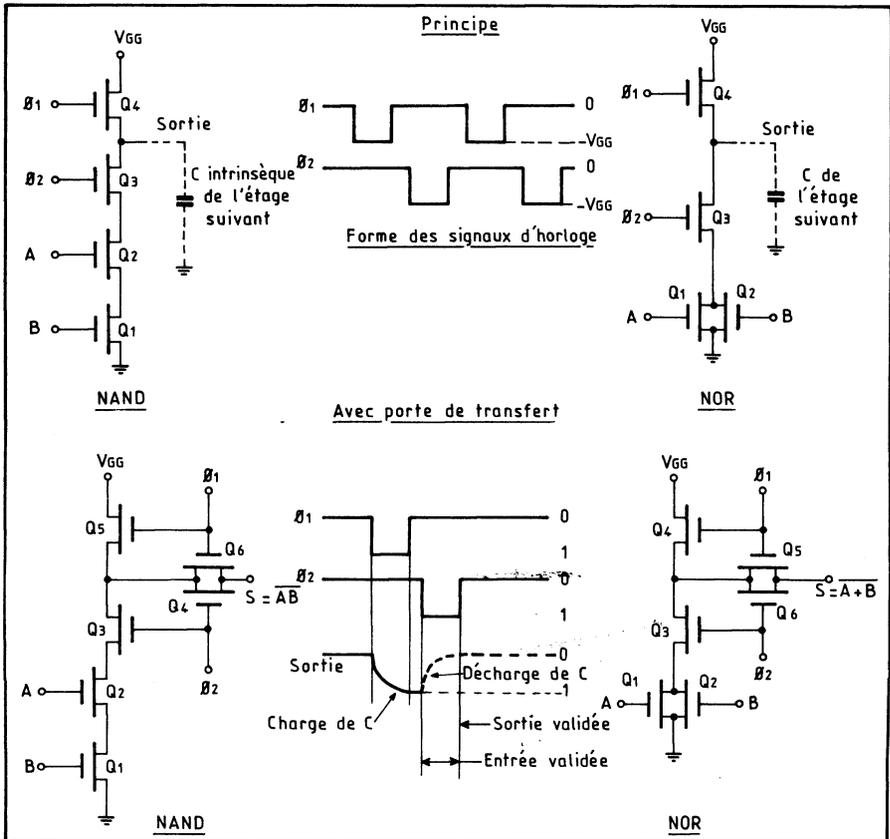


Fig. 12. — Deux structures de portes dynamiques non proportionnelles : NOR (a) et NAND (b), toutes deux à précharge.

Fig. 13. — Portes non proportionnelles à alimentation par les horloges.

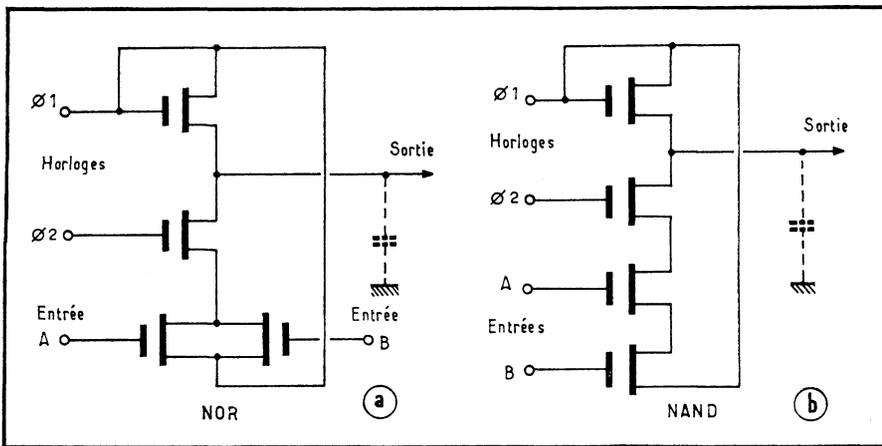
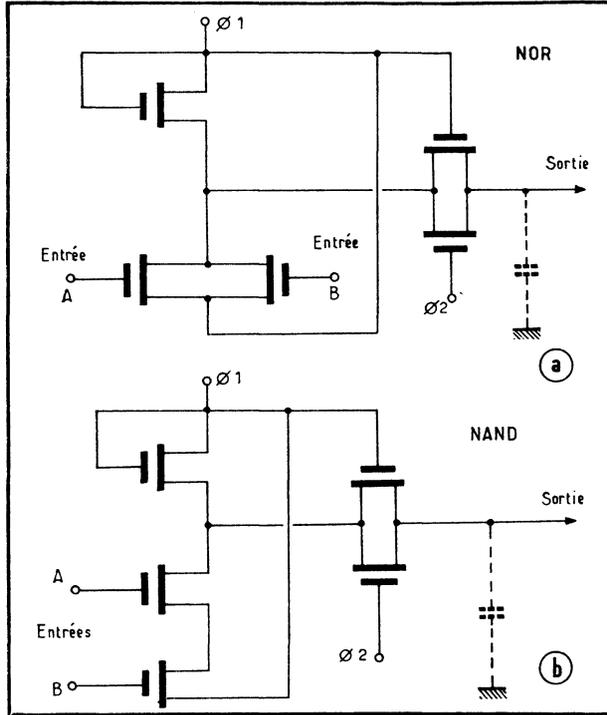


Fig. 14. — Autres structures de portes alimentées par les horloges.

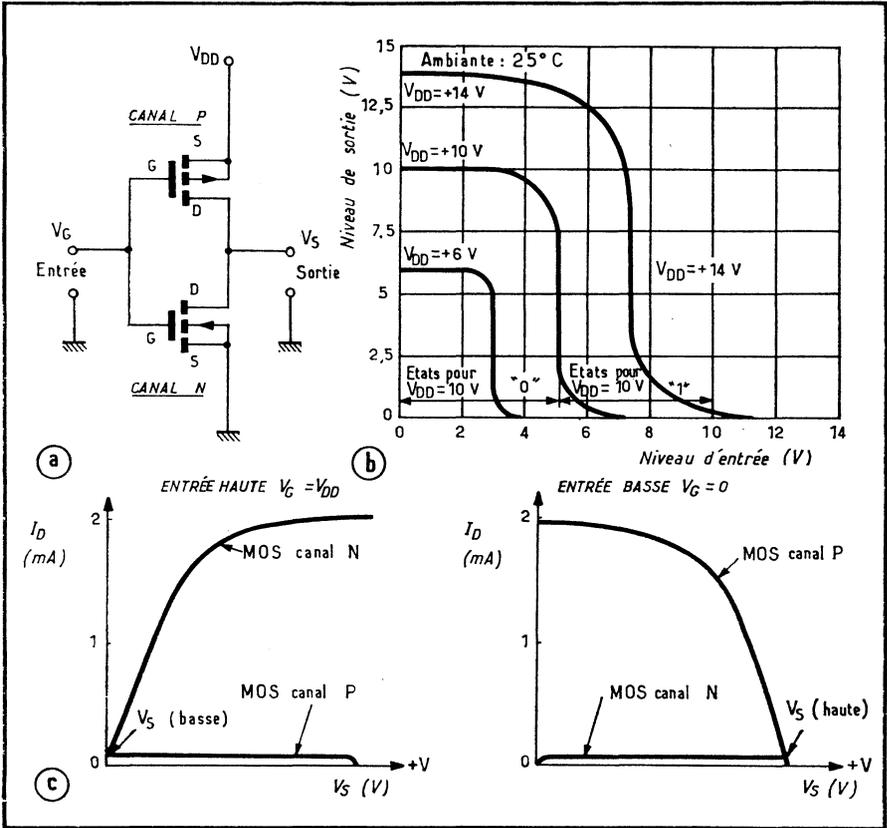


Fig. 15. — L'inverseur à MOS complémentaires (a) et ses courbes de transfert typiques (b). On notera la grande immunité au bruit, représentée ici pour $V_{DD} = 10V$ par les paliers horizontaux 0 et 1. En (c), courbes typiques relatives aux MOS individuels de la porte.

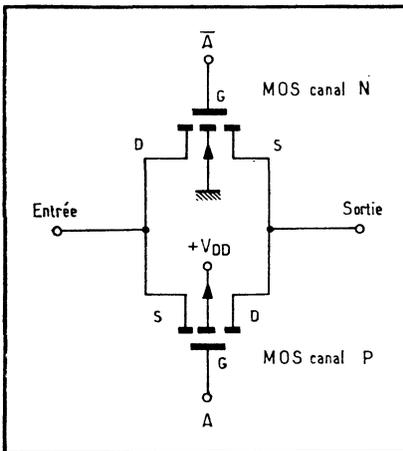


Fig. 16. — La porte de transfert à MOS complémentaires. Les deux transistors doivent être attaqués en opposition de phase.

La porte de transfert

La porte de transfert est ici constituée par deux MOS, l'un à canal P et l'autre à canal N (fig. 16). On applique les phases d'horloge en opposition, de façon que les deux transistors soient simultanément passants ou bloqués.

L'avantage, puisque cette porte est bilatérale, réside dans le fait qu'il y a toujours un MOS qui fonctionne avec charge de drain (alors que dans la logique monocanale à MOS canal P, par exemple, le MOS qui sert de transfert travaille soit à charge de drain, soit en source suiveuse, selon le sens de passage du courant).

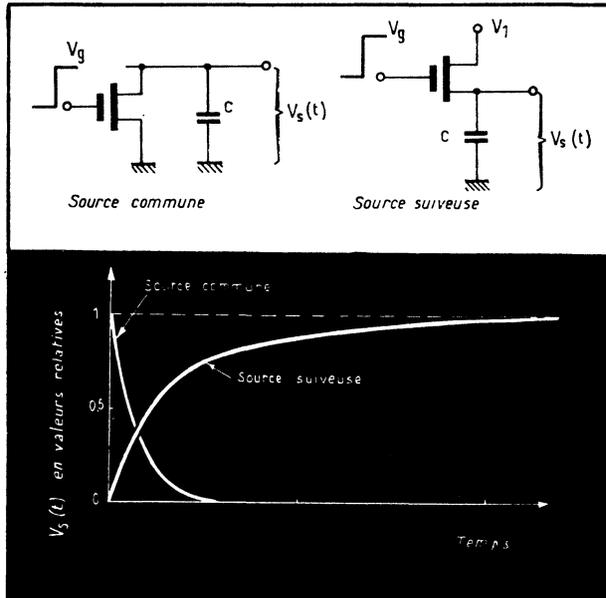


Fig. 17. — Ces courbes, relevées selon les deux montages dessinés ici, montrent que la commutation est plus rapide lorsque le MOS travaille en source commune, d'où l'avantage de la porte de transfert à CMOS.

On appréciera ce fait en se référant aux courbes donnant la tension de sortie des deux montages (à source suiveuse et à source commune, c'est-à-dire à charge de drain) de la figure 17 : celle correspondant à la charge dans le circuit de drain montre que la réponse à un signal en marche d'escalier est bien plus rapide (pour deux MOS du même type).

L'association d'un inverseur et d'une porte de transfert mène au commutateur représenté figure 18.

Porte NOR

Dans la porte NOR (fig. 19), deux MOS à canal N sont connectés en parallèle, et deux MOS à canal P sont montés en série.

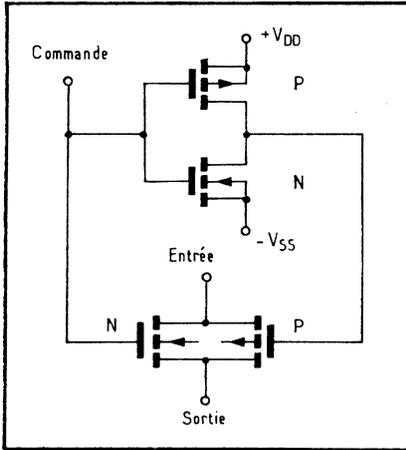


Fig. 18. — En associant un inverseur et une porte de transfert à CMOS, on crée un nouveau type de commutateur.

La sortie est à la masse lorsque l'entrée A ou l'entrée B (ou les deux) sont positives. A ce moment, en effet, les MOS correspondants sont :

- bloqué pour celui à canal P;
- conducteur pour celui à canal N.

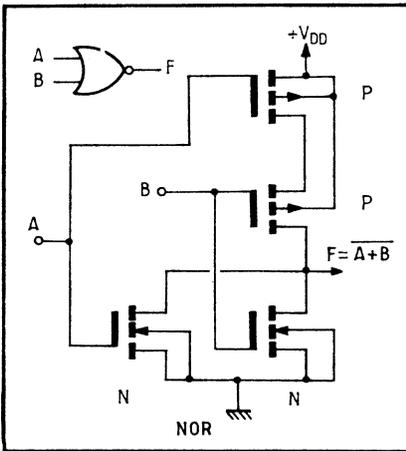


Fig. 19. — Porte NOR à deux entrées en CMOS.

On comprend qu'une tension nulle appliquée simultanément en A et B débloque les deux MOS à canal P, en série, et bloque ceux à canal N : la sortie passe à la tension positive.

On accroît le nombre des entrées en ajoutant, pour chaque entrée, un MOS canal P en série et un MOS canal N en parallèle.

Porte NAND

Une structure inversée réalise la fonction NAND (fig. 20) : deux MOS à canal P sont connectés en parallèle, et deux à canal N sont reliés en série.

La sortie passe à la masse lorsque les deux entrées A et B sont positives. Si l'une ou l'autre, ou les deux entrées sont à la masse, la chaîne série (MOS à canal N) est coupée tandis que la sortie est reliée à $+V_{DD}$ et devient positive.

A nouveau, on accroît le nombre des entrées en ajoutant à chaque fois un MOS canal P en parallèle et un MOS canal N en série.

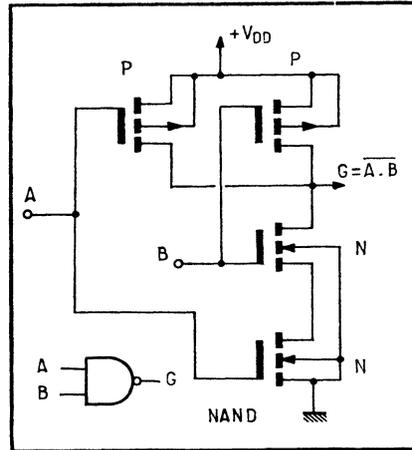


Fig. 20. — Porte NAND à deux entrées en CMOS.

6. LES INTERFACES POUR MOS MONOCANAUX

Par interfaces, on entend des circuits intermédiaires chargés de relier des familles MOS à d'autres familles, MOS, CMOS ou bipolaires, en transposant les

Tableau B. — Données pour l'interfaçage

NIVEAU LOGIQUE	MOS FONDAMENTAL canal P, haut seuil	DTL	TTL	ECL
0	-0,2V	0 à 0,5V	0 à 0,5V	-0,75V
1	-12 à -15 V	+5V	+3,5V	+1,55V
Alimentation	-13 et -27V	+5V	+5V	-5,2V

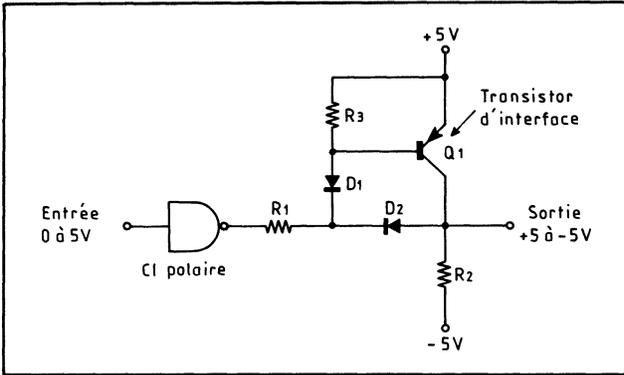


Fig. 21. — Interface simple bipolaire à MOS, 0 à 13 V.

niveaux lorsque besoin est, afin de les adapter l'une à l'autre. Ce problème n'est pas propre aux portes, avec lesquelles on va l'examiner, mais concerne aussi bien toutes les fonctions.

En fait, la logique la plus employée étant la bipolaire, TTL ou Schottky, puis les ECL et l'ancienne DTL, c'est avec ces familles qu'il faut compter (tableau B).

La TTL travaille sous +5 V et admet pour niveaux logiques :

- au zéro, 0,4 V max.
- au un, 3,6 V min.

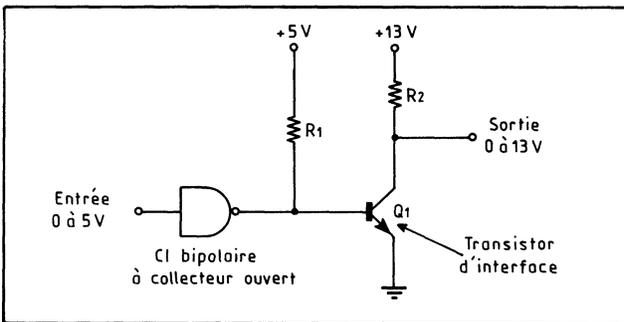


Fig. 22. — Interface bipolaire à MOS fournissant -5 et +5 V.

L'ECL est alimentée sous -5 ou -5,2 V et fournit des niveaux logiques (en logique positive) :

- au zéro, -1,55 V;
- au un, -0,75 V.

Or, le MOS fondamental à enrichissement, canal P, donc à haut seuil, est alimenté sous $V_{DD} = -13$ V et $V_{GG} = -27$ V; ses niveaux logiques sont :

- pour le zéro, -0,2 V;
- pour le un, -12 V en valeurs typiques.

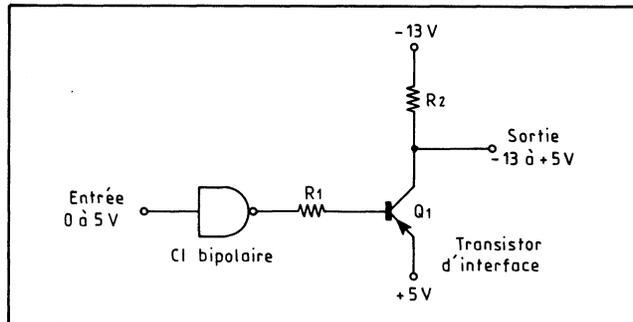
Les MOS ultérieurs, à bas seuil, canal P puis surtout canal N, ont été alimentés sous $+5\text{ V}$, ce qui harmonisait déjà l'alimentation, et occasionnellement les niveaux logiques.

Les interfaces constituent donc à chaque fois des cas d'espèces, pour lesquels quelques idées générales peuvent être dégagées. On verra ce qu'il en est pour les CMOS dans le chapitre qui leur est réservé.

Bipolaire à MOS

La translation bipolaire à MOS bas seuil et canal N est généralement directe. Avec le canal P, il faut veiller au passage de la *logique positive* à la *logique négative*.

Fig. 23. — Interface bipolaire à MOS fournissant -13 et $+5\text{ V}$.



L'inversion peut fort bien être assurée par un transistor, ce qui devient indispensable pour un MOS à haut seuil (fig. 21). Attaqué sur le gate, le MOS se présente comme une charge capacitive; la résistance R_1 est à supprimer lorsque le CI bipolaire est doté de sa charge.

Si le substrat du MOS est à $+5\text{ V}$ (ce cas se présente avec diverses technologies, celle à gate au silicium, par exemple), ce schéma devient celui de la figure 22. Les diodes évitent la saturation de Q_1 et réduisent le temps de blocage, mais ne sont pas impératives.

Ici, Q_1 est un NPN; si l'on utilise également un PNP dans le premier montage, on aboutit au schéma de la figure 23. Enfin, si le substrat du MOS est à la masse, on peut recourir au schéma de la figure 24.

Si le MOS est alimenté par $V_{SS} = +5\text{ V}$, la liaison peut encore être assurée directement, comme le montre la porte TTL de la figure 25. Le couplage direct avec le circuit TTL n'entraîne, pour ce dernier, aucun débit de courant en régime permanent quel que soit l'état logique. C'est le cas pour la plupart des circuits MOS pour lesquels on peut assimiler l'étage d'entrée à un inverseur simple. La charge est alors essentiellement capacitive (gate du MOS d'entrée), exception faite des courants de fuite négligeables. Dans ce type d'utilisation on peut remarquer qu'une porte TTL peut attaquer simultanément une autre porte TTL et un circuit MOS, ce qui peut contribuer à la simplification des systèmes dans lesquels les deux types de circuits coexistent.

La tension d'alimentation du MOS excédant 5 V , l'attaque pourrait encore être directe, mais ce type de liaison ne convient qu'à des TTL à sortie sur collecteur

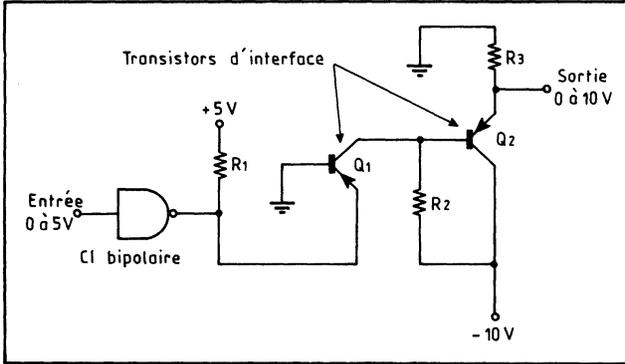


Fig. 24. — Interface bipolaire à MOS fournissant 0 à -10 V.

Fig. 25. — Liaison directe TTL à MOS (sous 5 V).

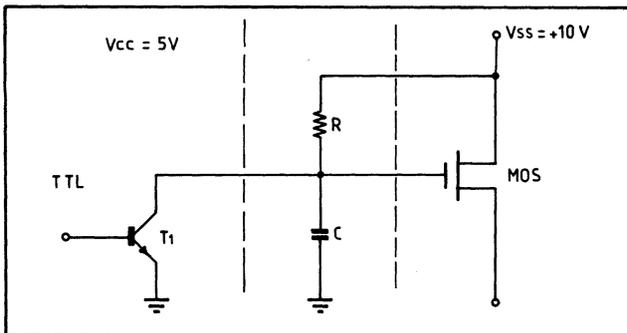
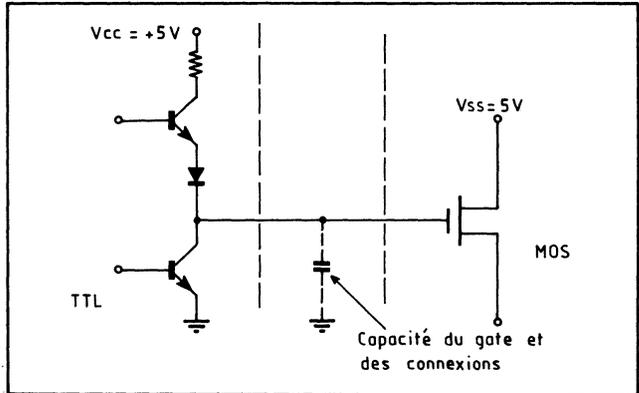


Fig. 26. — Utilisation d'une TTL à sortie à collecteur ouvert.

ouvert et non plus sur totem pôle, et dont, en outre, la tenue en tension de la sortie est supérieure à la tension V_{SS} (fig. 26). Or, la tenue en tension des sorties de la TTL standard est généralement inférieure aux 10 V donnés ici.

La résistance R intervient dans un double rôle de charge pour la TTL et de rappel pour le MOS; elle accroît l'excursion logique du signal à son entrée. A l'état zéro, il y a absorption de courant par le circuit TTL; avec les valeurs habituelles (0,4 V pour le zéro TTL et 16 mA max.), il faut donc que :

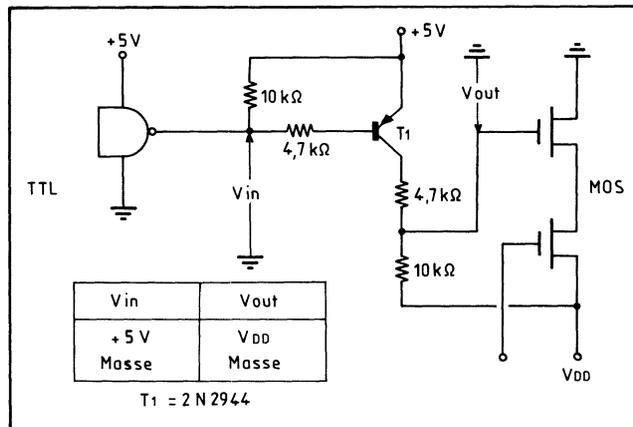
$$R > \frac{10 - 0,4}{16 \cdot 10^{-3}}$$

Soit

$$R > 600 \Omega$$

Le temps de montée du signal est principalement déterminé par la valeur de R et la capacité de charge C. Avec $R = 600 \Omega$ et $C = 50 \text{ pF}$, le temps de montée T

Fig. 27. — Interface active à PNP.



est 3 RC, soit 90 ns. Connaissant les capacités d'entrée des circuits MOS et le temps de montée maximal désiré, on obtient le nombre de circuits MOS qui peuvent être attaqués par une même porte TTL.

Enfin, il y a augmentation de la consommation globale du montage; on peut donc être amené à augmenter la valeur de la résistance R de manière à diminuer la consommation; si l'on désire garder un temps de montée du signal donné, il faudra également réduire le nombre de circuits MOS attaqués par une seule porte TTL.

Les figures 27 à 29, enfin, donnent trois autres exemples d'interfaces actives avec l'indication des excursions obtenues.

MOS à bipolaire (TTL)

Au niveau de l'adaptation MOS à TTL, le problème a deux aspects : assurer une compatibilité de niveau et absorber 1,6 mA à l'état « 0 » du TTL, le deuxième

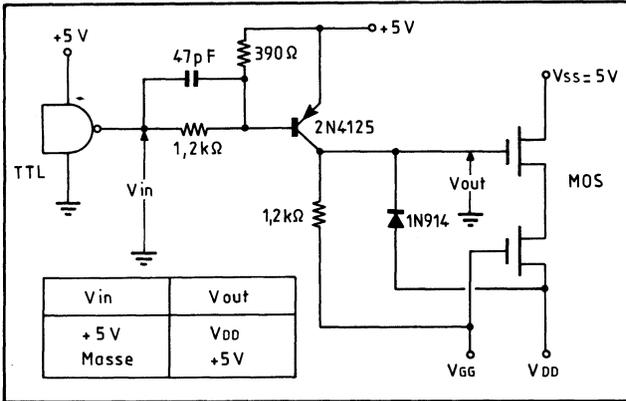


Fig. 28. — Interface active avec inversion des niveaux.

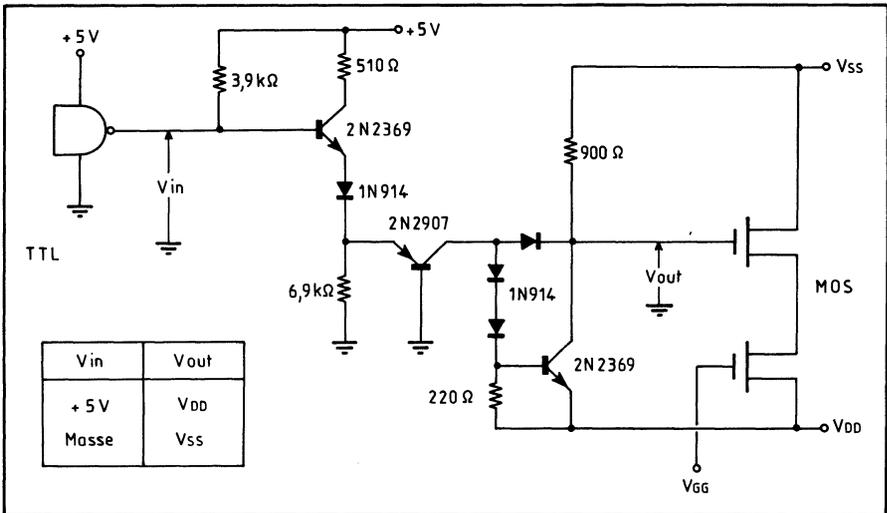


Fig. 29. — Interface à double transistor bipolaire.

Fig. 30. — Charge commune MOS-TTL.

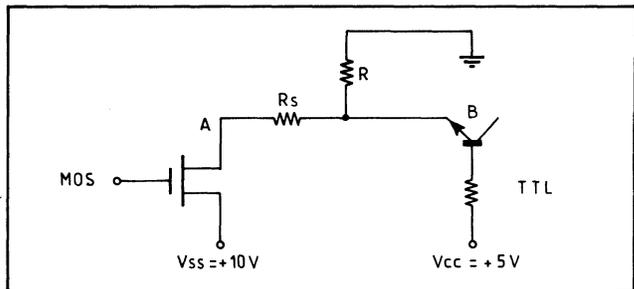


Fig. 31. — Un pont diviseur établit les niveaux souhaités.

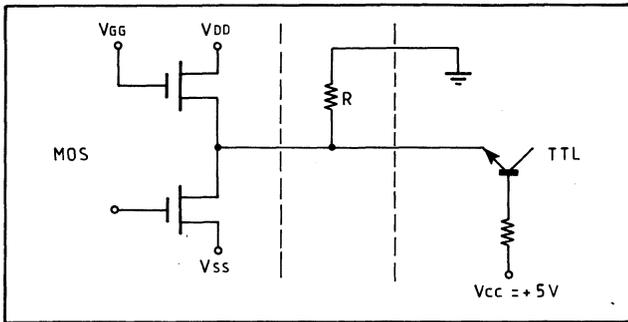
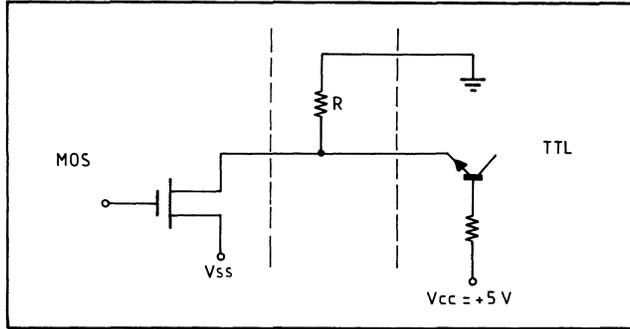
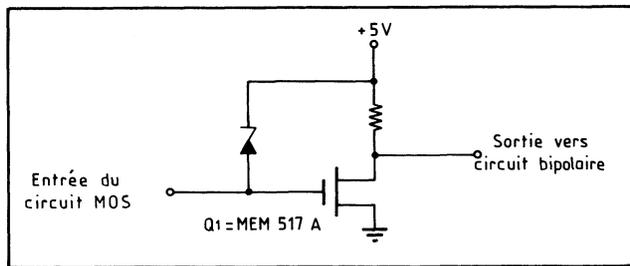


Fig. 32. — Le pont diviseur comprend ici un MOS et la résistance.

Fig. 33. — Interface simple MOS à bipolaire à un transistor MOS.



aspect étant de loin le plus contraignant. Les solutions apportées dépendent non seulement des conditions de polarisation du circuit MOS, mais également du type d'étage de sortie du circuit.

Le plus simple consiste à assurer une liaison directe, si le MOS est à drain ouvert, avec une résistance R faisant retour à la masse (fig. 30). Si $V_{SS} = +10\text{ V}$, une résistance série R_s , montée en diviseur avec R , ramène le niveau 1 à la valeur souhaitée (fig. 31). La dissipation est cependant importante.

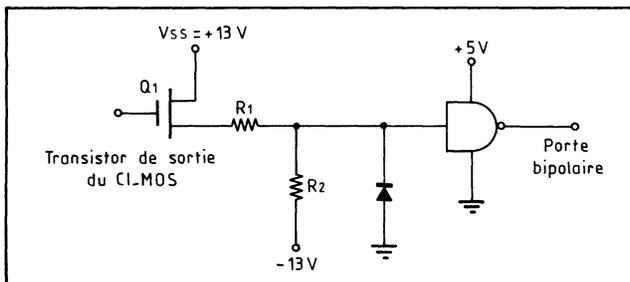
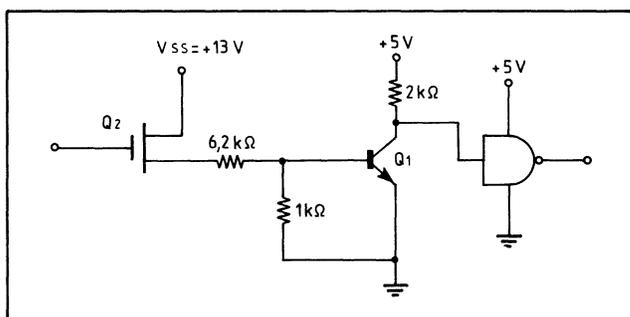


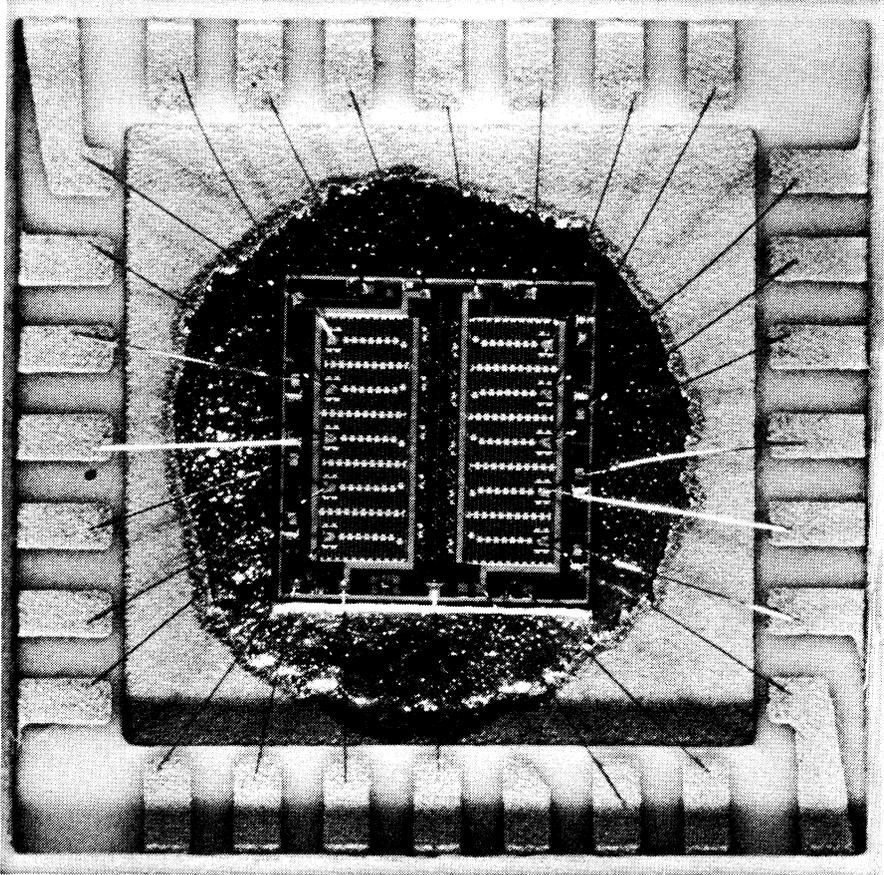
Fig. 34. — Interface à diviseur résistif MOS à bipolaire.

Fig. 35. — Interface MOS à bipolaire, à diviseur résistif avec transistor bipolaire de protection.



Si la sortie du circuit MOS comporte une charge MOS, une résistance R peut éventuellement suffire à aligner les niveaux (fig. 32); cette résistance doit être capable d'absorber 1,6 mA sous 0,4 V au niveau 0, soit 250 Ω . Elle pourrait être ramenée à une tension négative, à la condition de disposer d'une diode de blocage à l'entrée du TTL. Les puissances limites devront être respectées dans tous les cas.

Une autre interface simple MOS à bipolaire avec transistor actif MOS intermédiaire est indiquée figure 33. Tant que Q_1 est bloqué, la sortie est à +5 V; sinon, la sortie est à la masse. Un diviseur résistif permet une liaison directe (fig. 34) à partir des +13 et -13 V, la diode interdisant à l'entrée de la porte bipolaire de devenir négative par rapport à la masse. Avec des tensions de +13 V et masse, l'interface peut se ramener à un NPN, comme le montre la figure 35.



Montage d'une puce de circuit intégré. On distingue nettement les fils de connexions qui relient la puce de silicium aux pattes métallisées du boîtier. (Doc. TN.)

CHAPITRE XI

LES FLIP-FLOPS

1. LES PRINCIPALES BASCULES

L'arsenal des bascules, ou flip-flops, comprend 4 types de base, le RS, le JK, le D et le T, dont les tables d'états sont comparées dans le tableau A.

Le RS et le JK disposent tous deux de deux entrées, dont les combinaisons déterminent l'état des sorties. Celles-ci sont quasiment toujours doubles et fournissent l'information et son inverse; dans le tableau, on n'a cependant retenu que la situation de la sortie « vraie » notée Q. Les temps sont notés Q_{n+1} pour l'état suivant, après exécution de l'ordre appliqué aux entrées ou après l'impulsion d'horloge, et Q_n pour l'état précédant. Le point d'interrogation indique un état indéterminé.

Les flip-flops D et T sont à entrée unique; le D est généralement synchrone, alors que les RS et JK peuvent être asynchrones (schéma de base) ou synchrones. Le D existe, de plus, en deux variantes : D à *commutation sur le front avant* et D à *verrouillage*. On retrouvera cette même notion de verrouillage pour des flip-flops dit « latchés » (de l'Américain « *latch flip-flop* ») avec les autres types de bascule, commutant sur le front arrière (descendant).

Tableau A. — Les 4 flip-flops fondamentaux

Entrées pour RS ou JK		Sortie Q_{n+1} (état suivant), avec Q_n = état précédent				Entrée pour
R ou J	S ou K	RS	JK	D	T	D ou T
0	0	Q_n	Q_n			—
0	1	1	1	1	$\overline{Q_n}$	1
1	0	0	0	0	Q_n	0
1	1	?	$\overline{Q_n}$			—

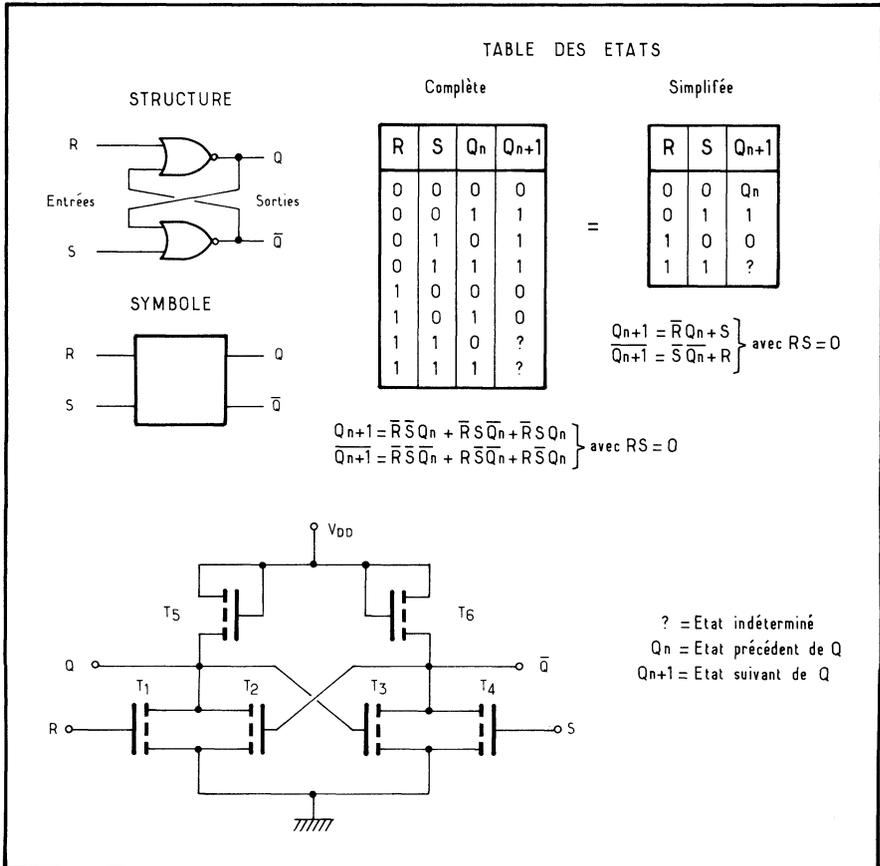


Fig. 1. — Le RS fondamental.

Parmi les variantes figurent essentiellement les *maître-esclave*, qui s'adressent aux quatre types de base.

Enfin, on trouvera des flip-flops *statiques* ou *dynamiques*, et des montages à MOS monocanaux ou complémentaires. Ces derniers seront examinés séparément ici.

2. LE FLIP-FLOP RS

Le flip-flop le plus simple est le RS, réalisé à l'aide de deux portes NOR rétro-couplées (fig. 1). Les deux entrées sont :

- R, de *reset*, ou remise à zéro, car elle force la sortie Q à 0 et par conséquent, \bar{Q} à 1;
- S, de *set*, ou mise à 1, car elle force la sortie Q à 1, et par conséquent, \bar{Q} à 0.

D'après le schéma, on voit en effet que si l'on fait $R = 0$ et $S = 1$, le transistor T_4 conduit et \bar{Q} passe à 0; simultanément, T_2 se bloque et comme T_1 est

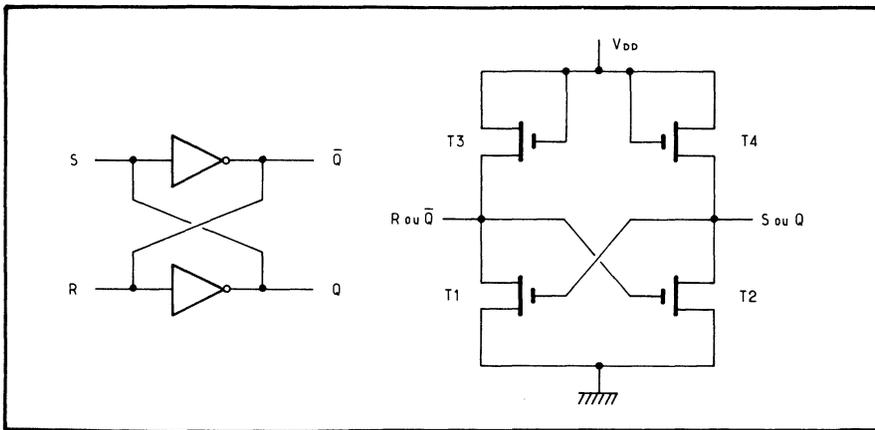


Fig. 2. — RS simplifié à deux inverseurs.

également bloqué, Q passe à 1. Les ordres d'entrée R et S disparaissant, les états de sortie sont verrouillés par la bascule et maintenus « en mémoire ». Quatre situations sont possibles :

1. $R = 0$ et $S = 1$, comme on l'a vu. On a : $Q = 1$ et $\bar{Q} = 0$;
2. $R = 1$ et $S = 0$. C'est l'inverse de la précédente; elle provoque $Q = 0$ et $\bar{Q} = 1$;
3. $R = 0$ et $S = 0$: dans ce cas, la situation précédent cet état est maintenue et il n'y a pas de changement de l'état des sorties, ce qui s'écrit : $Q_{n+1} = Q_n$;
4. $R = 1$ et $S = 1$: c'est la situation inconfortable pour le RS, car elle provoque une situation imprévisible à la sortie, quand ce n'est pas une instabilité ou une oscillation incontrôlée. Elle est donc interdite, ce qui s'exprime logiquement

par la condition $RS = 0$, pour que ce type de flip-flop puisse fonctionner ($RS = 1$ interdit).

Le RS est réalisé, en MOS, à l'aide de portes NOR car les NAND (utilisées en bipolaires), se prêtent moins bien, ici, à des structures de bascules.

RS à inverseurs

Pour certaines applications, la structure à portes NOR n'est pas obligatoire et l'on peut simplifier le RS en ne recourant qu'à deux inverseurs (fig. 2). En effet, on conserve le même principe de verrouillage mais les entrées et sorties ne sont plus différenciées; d'autre part, les situations $R = S$ sont interdites, aussi bien $R = 0$ et $S = 0$ que $R = 1$ et $S = 1$.

RS à entrées multiples

Les conditions d'entrées peuvent être multipliées, ce qui se fait couramment, en utilisant soit des NOR à plusieurs entrées, soit toutes autres structures. La figure 3 représente quelques symboles de RS à entrées multiples.

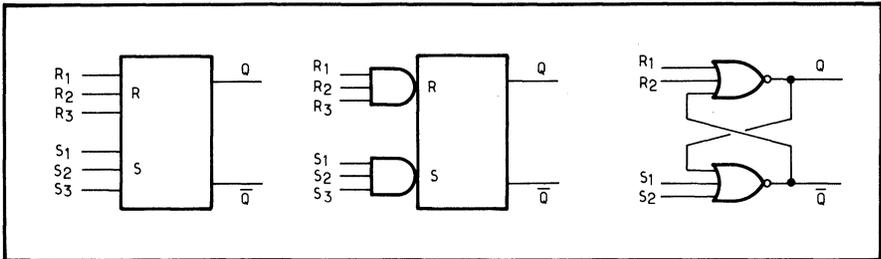


Fig. 3. — Trois exemples de dessin de RS à entrées multiples.

Le RS synchrone

La transformation du RS asynchrone en RS synchrone est réalisée en affectant un MOS supplémentaire, en ET, au schéma de base (fig. 4). En fait, ce sont deux paires AND/NOR que l'on associe dans une structure dont le fonctionnement est évident.

On voit que, tant que l'horloge est *haute*, les informations d'entrée ne doivent pas changer car sinon, les sorties changeraient également d'état en fonction de ces ordres : *elles ne sont verrouillées que sur le front descendant de l'horloge*. C'est pourquoi l'on désigne encore cette bascule par *RS à verrouillage* (le *RS latch* des Américains).

Une autre méthode consiste à réaliser un montage où les MOS d'entrée sont reliés à la masse via un transistor commandé par la phase d'horloge (fig. 5). On économise ainsi un MOS. Dans les deux cas, c'est une seule phase d'horloge qui sert de commande.

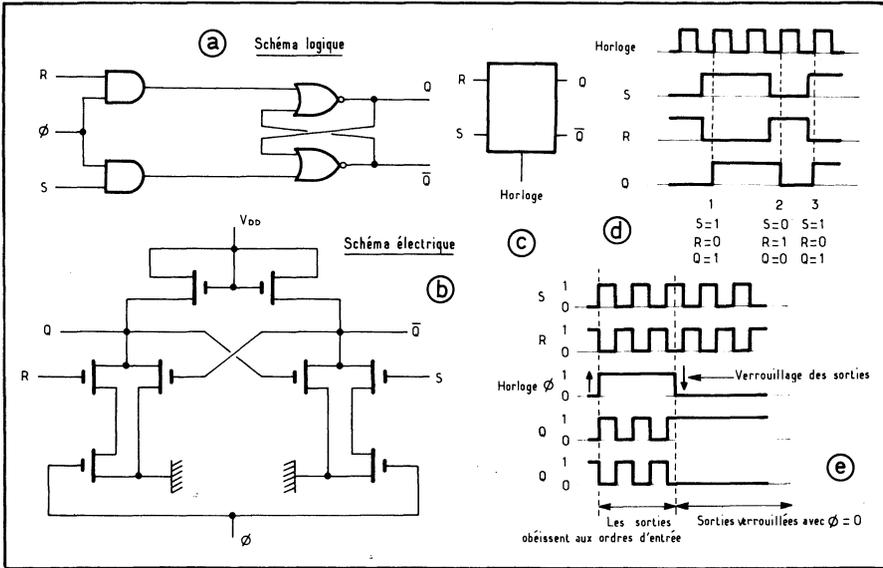
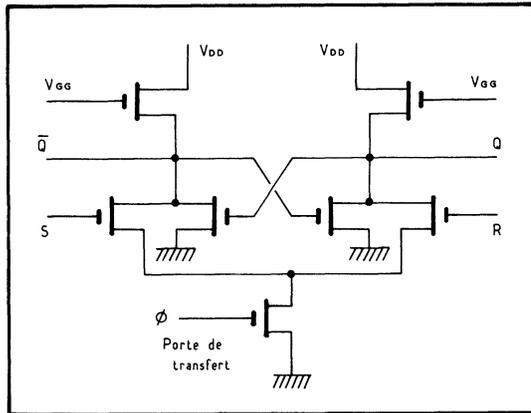


Fig. 4. — Le RS synchrone.

Fig. 5. — Autre version du RS synchrone.



RS à sorties trois états

Une autre structure, qui pourrait s'apparenter à un mode synchrone, est réalisée avec l'utilisation de portes de transfert à la sortie (fig. 6). Ces portes sont conductrices lorsqu'un signal d'autorisation leur est appliqué; sinon, elles isolent la sortie de la bascule. Ainsi aboutit-on à un circuit à trois états :

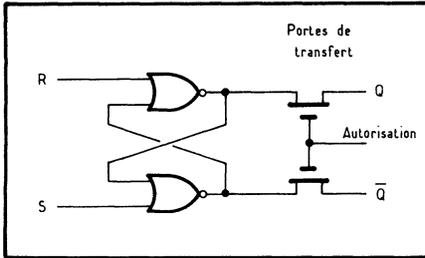


Fig. 6. — Le RS à sorties 3 états.

- *état bas*, la porte de transfert étant activée et $Q = 0$;
- *état haut*, la porte de transfert étant activée et $Q = 1$;
- *état à haute impédance* (sortie « en l'air »), la porte n'étant pas activée, pour $Q = 0$ ou 1 indifféremment.

RS en MOS complémentaires

La cellule RS en CMOS fait appel à des MOS canal P et canal N, selon la structure de la figure 7; on retrouve là les deux portes NOR rétroconnectées :

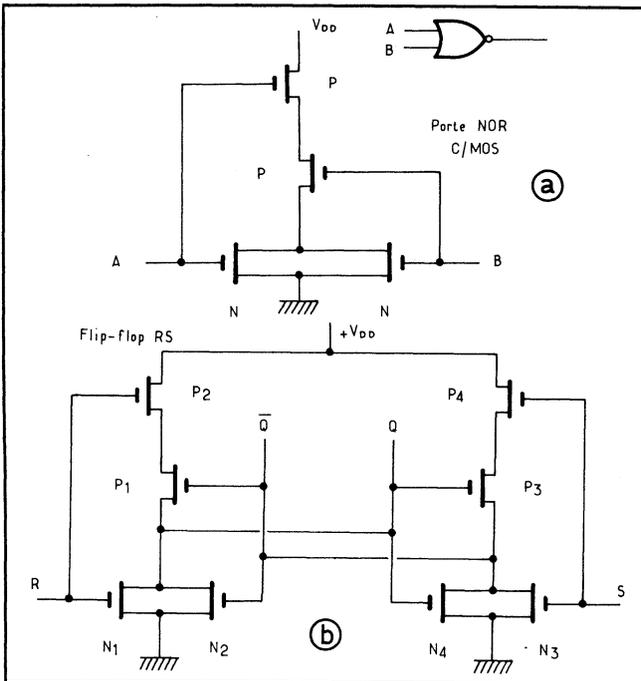


Fig. 7. — Le RS en CMOS se compose de deux portes NOR telles que celle dessinée en a. La structure RS est donnée en b.

- Pour $R = 1$ et $S = 0$, les transistors N_1 (canal N) et P_4 (canal P) sont conducteurs alors que P_2 et N_3 sont bloqués. La sortie Q passe donc à 0, ce qui débloque P_3 ; par conséquent, \bar{Q} passe à 1, bloquant P_1 .
- Si $S = 1$ et $R = 0$, on aboutit à la situation inverse, soit $Q = 1$ et $\bar{Q} = 0$.
- Pour $R = S = 0$, on voit que N_1 et N_3 sont bloqués alors que P_2 et P_4 sont conducteurs. Si l'on avait $Q_n = 1$ et $\bar{Q}_n = 0$, les transistors N_4 et P_1 conducteurs confirmeraient cette situation en faisant $Q_{n+1} = 1$ et $\bar{Q}_{n+1} = 0$. L'inverse $Q_n = 0$ et $\bar{Q}_n = 1$ serait confirmé de la même façon.
- Enfin, pour $R = S = 1$, ce sont N_1 et N_3 qui deviennent passants alors que P_2 et P_4 restent bloqués; les deux sorties passent à zéro!

Lors de certaines applications aux mémoires, par exemple, la structure élémentaire à deux inverseurs peut être retenue (fig. 8).

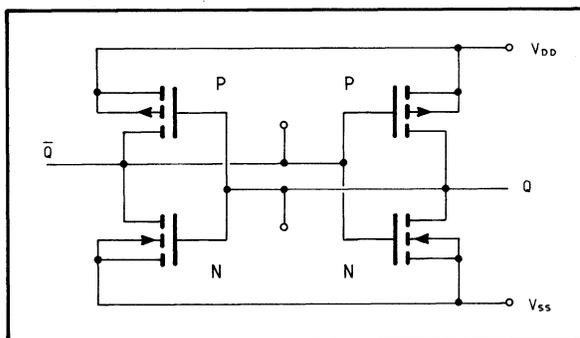


Fig. 8. — Version simplifiée du RS à deux inverseurs CMOS.

L'exemple d'un RS réel et complet, doté d'une porte de transfert sur la sortie unique, est donné en CMOS figure 9. Il s'agit du circuit *CD 4043*, qui est en réalité une quadruple bascule RS (4 bascules identiques dans un unique boîtier). On a dessiné, également, les circuits de protection des entrées, ainsi qu'un circuit d'auto-risisation qui, appliqué à la porte de transfert, en fait un montage à trois états.

3. LES RS DYNAMIQUES

En recourant à des portes dynamiques, on constitue des flip-flops dynamiques dont la figure 10 donne un exemple, à une seule phase d'horloge.

- Supposons que $R = 0$ et $S = 1$; le MOS noté M_8 est conducteur, donc Q passe à 0, tandis que M_5 est bloqué, ce qui entraîne $Q = 1$. L'application de la phase d'horloge Φ vient confirmer l'état de la bascule en transmettant 1 à M_1 et 0 à M_2 , seconds éléments des OU d'entrée. Comme on le voit, les capacités de structure C_{GS} mémorisent les états.

- A l'inverse, $R = 1$ et $S = 0$ mènent à $Q = 0$ et $\bar{Q} = 1$. La situation $R = S = 0$ maintient la situation antérieure, alors que $R = S = 1$ fournit un état indéterminé des sorties; c'est à nouveau une situation interdite.

En fait, les capacités de structure n'interviennent que tant que l'horloge est basse; dès qu'elle devient haute, le fonctionnement est du type statique. Aussi ne

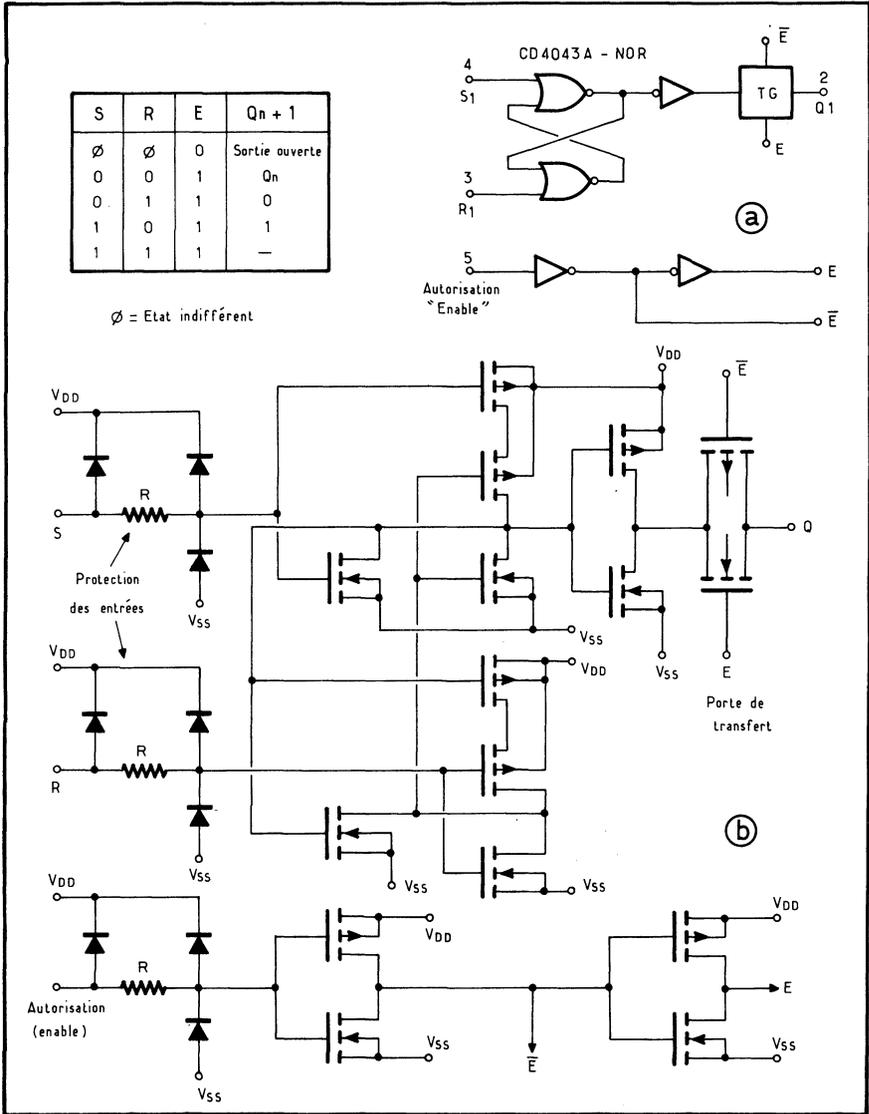
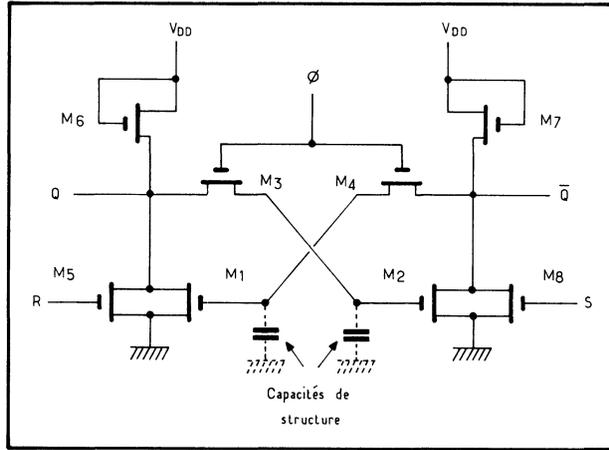


Fig. 9. — Le RS type CD 4043, à portes NOR. Ce CI est, en réalité, un quadruple RS dont on ne voit ici qu'une bascule : synoptique (a) et schéma (b)

sera-t-on pas surpris de voir qualifier ce type de montage de dynamique, de statique, ou de semi-dynamique ou semi-statique, selon les auteurs.

On peut ramener une telle structure à deux inverseurs si l'on consent à introduire R et S directement sur les gates de M₁ et M₂ (fig. 11). Il est cependant nécessaire d'ajouter deux MOS supplémentaires, commandés par une seconde phase

Fig. 10. — Le RS dynamique.



d'horloge, pour permettre le fonctionnement normal de la bascule lors de changements d'états : sinon, $Q = 1$ pour $S = 0$, par exemple, mènerait à un court-circuit pendant la phase d'horloge.

Le fonctionnement est illustré par les chronogrammes de la figure. Supposons que $R = 0$ et $S = 1$; pendant le temps t_1 , les phases Φ_1 et Φ_2 sont à zéro et le système est verrouillé. En t_2 , Φ_1 passe à 1; par conséquent, R et S sont appliqués à M₁ qui se bloque et M₂ qui devient conducteur; donc, Q passe à 1 et \bar{Q} à 0.

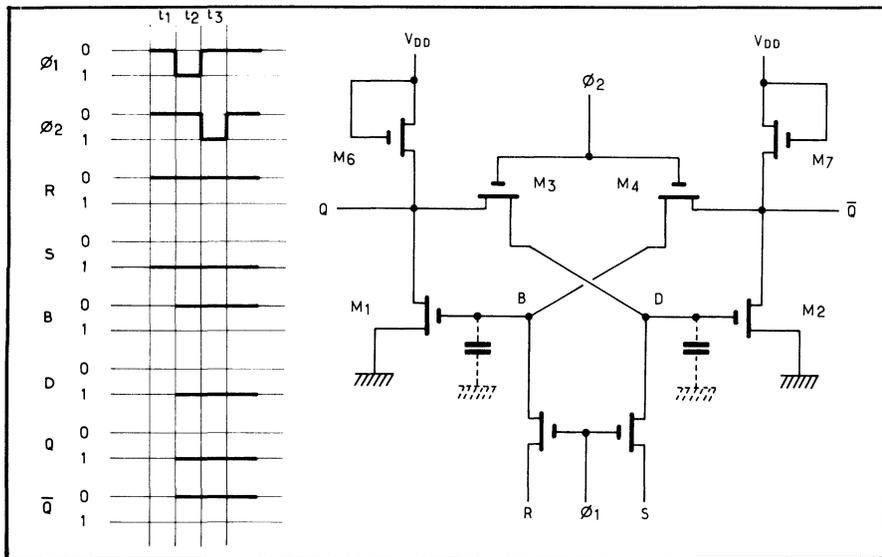


Fig. 11. — Structure simplifiée du RS dynamique.

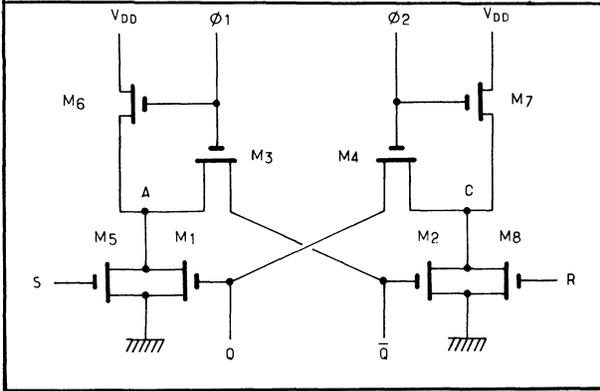


Fig. 12. — Autre variante du RS dynamique.

En t_3 , la phase Φ_1 revient à 0 alors que Φ_2 passe à 1 : la mise en conduction de M_4 et M_3 , en reportant les niveaux de sortie aux gates de M_1 et M_2 , vient confirmer l'état acquis. Par contre, $R = S = 0$ ou $R = S = 1$ ne sont pas admissibles ici ; ils apparentent cette structure à une bascule T, qu'on verra plus loin.

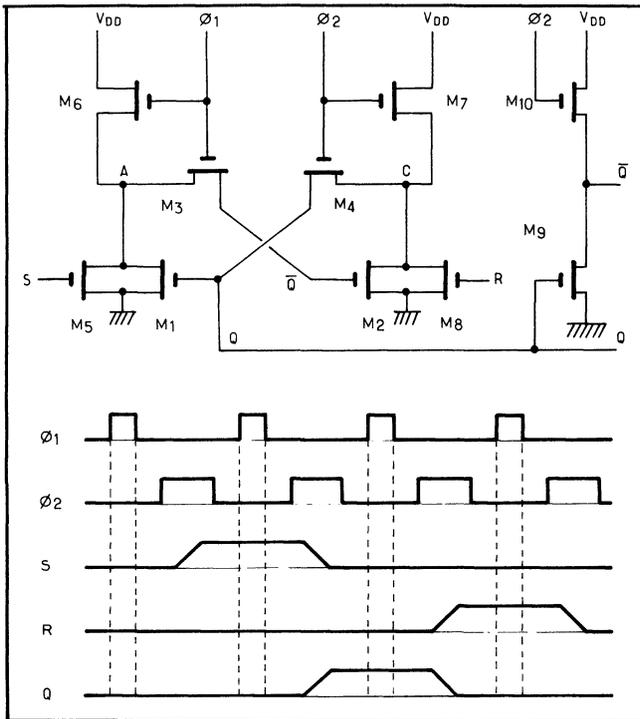


Fig. 13. — Il peut être intéressant d'utiliser une porte inverseuse à la suite du montage flip-flop pour disposer d'un signal complémentaire.

On retrouve également deux phases d'horloge avec des portes dynamiques dans la figure 12. On ne peut plus alors placer les sorties Q comme précédemment puisque les portes NOR peuvent chacune se trouver totalement bloquées, les charges M_7 ou M_6 se trouvant non conductrices en l'absence de signaux d'horloge. Elles sont donc prélevées directement aux gates de M_1 et M_2 .

Comme il peut être gênant de charger ainsi les capacités de structure, on peut faire suivre les sorties par des portes, comme le montre la figure 13 où apparaissent les chronogrammes typiques.

Avec de telles structures, on se trouve, en réalité à la limite avec les cellules de registre à décalage, comme on le verra ultérieurement. C'est ce dont témoigne encore le schéma suivant; c'est celui d'une cellule dynamique plus complexe proposée par la firme américaine *Four Phase System* (fig. 14); elle est constituée par deux étages inverseurs, avec précharge, et exige quatre phases d'horloge.

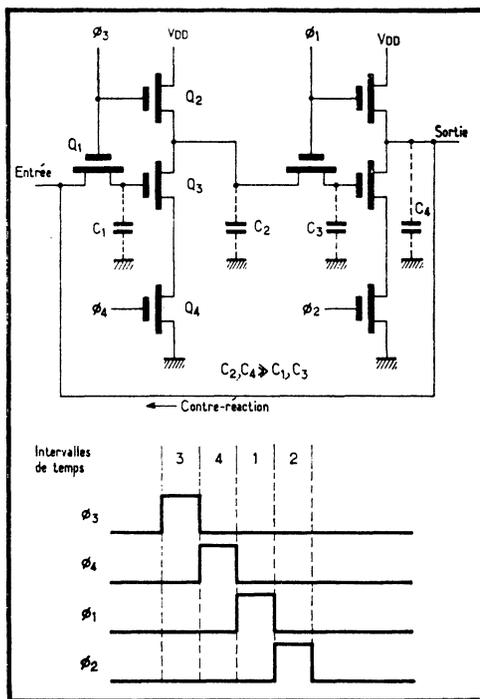


Fig. 14. — Flip-flop dynamique quadriphase, à précharge.

Avec la phase Φ_3 , le transistor Q_1 est débloquenté et la capacité de structure C_1 est chargée au niveau du signal d'entrée (0 ou 1). Simultanément, Q_2 conduit également et charge C_2 : cette capacité gate-substrat est préchargée au niveau du V_{DD} .

Puis survient la phase Φ_4 , qui débloquenté Q_4 ; si la charge de C_1 est au niveau 1, le transistor Q_3 entre en conduction et décharge C_2 à la masse. Dans le cas contraire (C_1 au niveau 0), C_2 reste chargé.

La phase Φ_1 sert à transmettre le niveau de C_2 à la seconde cellule, où Φ_1 et Φ_2 agissent de la même façon que Φ_3 et Φ_4 en inversant le signal, c'est-à-dire en le restaurant dans sa phase primitive.

En raison du fait que la sortie est ensuite ramenée à l'entrée, l'information est verrouillée; on a ainsi transformé ce registre à décalage à 1 bit en élément binaire, en un flip-flop dynamique qui mémorise une information.

4. LA BASCULE JK

Le flip-flop RS comporte une combinaison interdite à l'entrée parce qu'elle mène à une situation indéterminée à la sortie.

Or, il devrait être possible d'utiliser les informations présentes en Q et \bar{Q} pour éliminer, par aiguillage, cet état indéterminé. C'est là ce qui est fait avec le flip-flop JK, ainsi dénommé parce que ces lettres étaient encore disponibles pour en désigner les entrées.

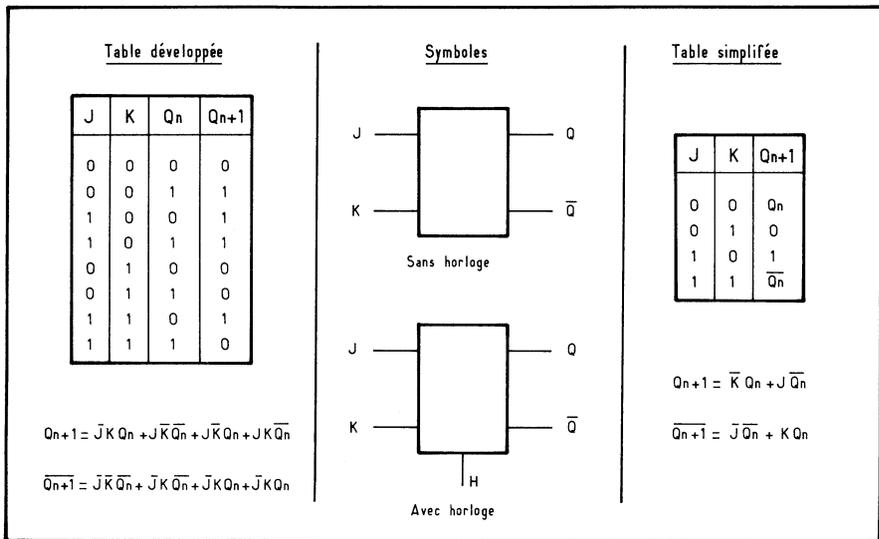


Fig. 15. — Caractéristiques du JK.

Le symbole du JK est donné figure 15 avec ses tables développée et simplifiée des états. On constate que toutes les situations possibles aux entrées déterminent des états précis aux sorties. Ainsi :

- l'entrée J se comporte comme une entrée de mise à 1 (« set »);
- l'entrée K se comporte comme une entrée de mise à 0 (« reset », ou « clear »);
- deux niveaux 0 aux entrées laissent la bascule dans le même état;
- deux niveaux 1 aux entrées provoquent l'inversion des états précédents.

On va examiner son fonctionnement à l'aide du schéma de principe de la figure 16, très proche des schémas réels à MOS. On remarquera la connexion de rebouclage qui part de la sortie Q et ramène l'information aux entrées NOR et AND. Dans la figure ont été portées diverses situations, et l'on passe de l'une à

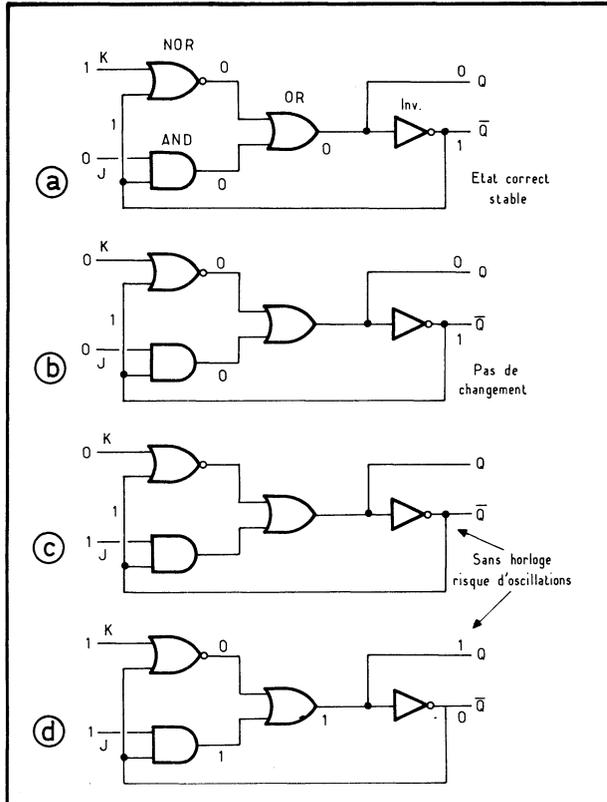
l'autre quant aux états; enfin les tables de vérité des portes sont rappelées pour mémoire.

- En *a*, on fait $J = 0$ et $K = 1$. La sortie du NOR est donc forcément 0, et celle du AND forcément 0 aussi. Donc, le OR fournit un 0 en Q , et un 1 en \bar{Q} .
- De là, on passe en *b* où $J = K = 0$, le 1 de \bar{Q} étant réappliqué aux entrées. On peut aisément vérifier que les sorties ne changent pas. Mais si l'on fait ensuite $K = 0$ et $J = 1$, en *c*, ou $K = J = 1$, en *d*, on s'aperçoit vite que le système ne fonctionne plus et risque d'entrer en oscillations.

Fig. 16. — Principe de fonctionnement du JK, pour plusieurs conditions d'entrée.

Pour en suivre le fonctionnement, on se rappellera que :

A	B	AND	OR	NOR
0	0	0	0	1
0	1	0	1	0
1	0	0	1	0
1	1	1	1	0



Il convient donc d'introduire un élément de retard en série avec \bar{Q} pour reporter cette information à l'entrée; le principe en est dessiné figure 17, où les entrées sont J et K. Avec $V_{DD} = 1$ et la masse = 0, on voit que, pour :

1. $J = 0$ et $\overline{K} = 1$, soit $\overline{K} = 0$.
 - Si $Q_n = 0$ et $\overline{Q}_n = 1$, le point A = 1 et $\overline{Q}_{n+1} = 1$, donc les sorties ne changent pas.
 - Si $Q_n = 1$ et $\overline{Q}_n = 0$, la chaîne reste coupée entre A et \overline{Q}_{n+1} , et l'on obtient encore $Q_{n+1} = 1$ et $Q_{n+1} = 0$.
2. $J = 1$ et $K = 0$, soit $K = 1$, la sortie $\overline{Q}_{n+1} = A = 0$, puisque la chaîne est conductrice, et ce quels que soient Q_n et \overline{Q}_n . Donc, $Q_{n+1} = 1$.
3. $J = K = 0$, donc $J = 0$ et $K = 1$: tout dépend ici de Q_n et \overline{Q}_n :
 - Pour $Q_n = 0$ et $\overline{Q}_n = 1$, on a $A = 1$ et donc $\overline{Q}_{n+1} = 1$, par conséquent $Q_{n+1} = 0$. Les sorties n'ont pas changé.
 - Pour $Q_n = 1$ et $\overline{Q}_n = 0$, la chaîne est conductrice et $A = Q_{n+1} = 0$, donc $Q_{n+1} = 1$. A nouveau, les sorties n'ont pas changé.

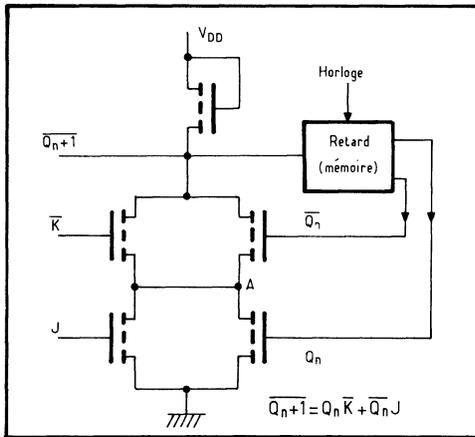


Fig. 17. — Structure de principe du JK.

4. $J = K = 1$, donc $J = 1$ et $\overline{K} = 0$, le point A = 0, forcément, donc \overline{Q}_{n+1} dépend de \overline{Q}_n :

- Si $Q_n = 0$, on obtient $\overline{Q}_{n+1} = 1$.
- Si $\overline{Q}_n = 1$, on obtient $\overline{Q}_{n+1} = 0$.

Il y a donc eu inversion des sorties, et l'on retrouve ainsi la table de vérité du JK.

L'élément retard, chargé de mémoriser temporairement l'information, est bien évidemment une capacité de structure de MOS; cette information est ramenée à l'entrée après un retard imposé par les deux phases d'horloge. On aboutit ainsi au schéma de la figure 18, où l'on retrouvera la première structure évoquée au début de ce paragraphe où ne figuraient pas les phases d'horloge.

Une autre structure, conforme à celle de principe avec phases d'horloge, est donnée figure 19. Que l'on raisonne en logique positive ou négative, on voit que Φ_1 enregistre l'information présente aux entrées tandis que Φ_2 fait la décision aux sorties.

Pour préciser ce que l'on entend par *temps de propagation*, on a rappelé, figure 20, les principales définitions de principe, à partir d'une unique phase d'horloge (avec 2 phases, c'est la période d'horloge qui entre en ligne de compte) :

- le temps de montée du signal d'horloge est appelé t_r (de *rise time*); il se mesure de 10 à 90 % de l'amplitude totale;

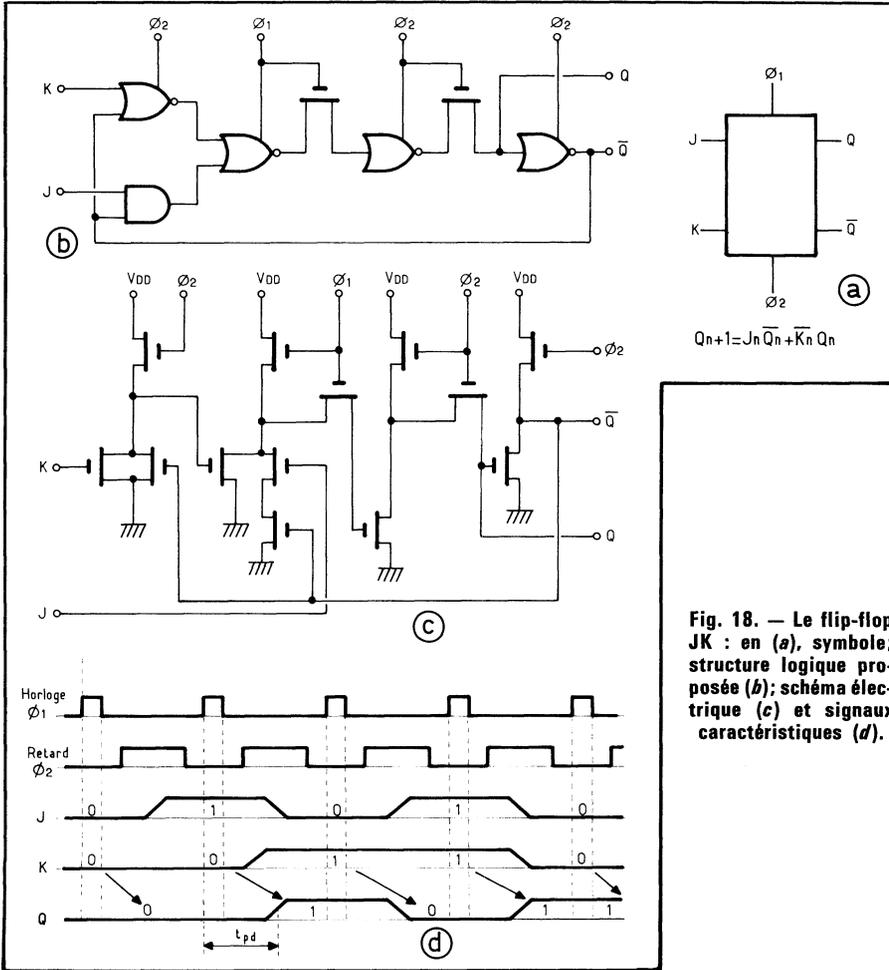


Fig. 18. — Le flip-flop JK : en (a), symbole; structure logique proposée (b); schéma électrique (c) et signaux caractéristiques (d).

- l'impulsion d'horloge doit disposer d'une largeur t_p suffisante; elle se mesure, par convention à 50 %;
- les informations, sur les entrées J ou K, doivent être présentes un certain temps avant le signal d'horloge : c'est le *temps de préconditionnement*, ou « *set up time* »;

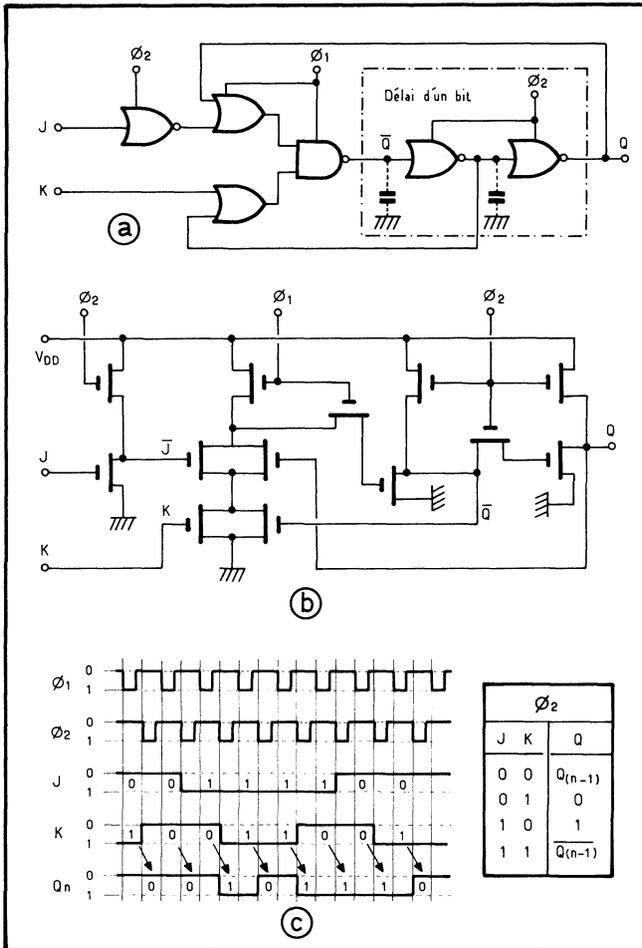
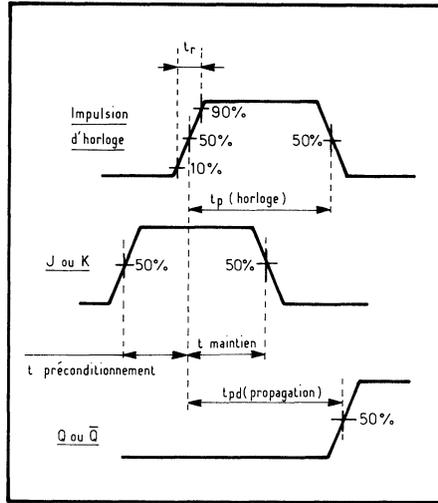


Fig. 19. — Autre exemple de JK, utilisant en particulier une porte complexe NAND/NOR : structure logique (a), schéma électrique (b), et signaux obtenus (c).

- ces impulsions doivent être maintenues pendant un certain temps après le démarrage de l'impulsion d'horloge : c'est le *temps de maintien*, ou « *hold time* », noté $t(\text{hold})$.
- le retard séparant la montée de la première impulsion d'horloge et le basculement du JK est le *temps de propagation* (« *propagation delay time* »), noté t_{pd} .

Fig. 20. — Signaux devant être appliqués aux entrées (J ou K) et à l'horloge et signal résultant à la sortie pour un flip-flop JK.



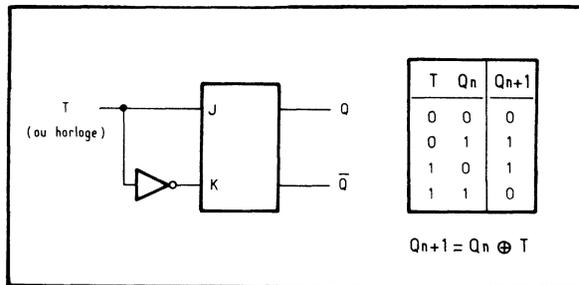
5. LE FLIP-FLOP T

Le flip-flop T (de *trigger*, ou *toggle*), bascule lors de chaque nouvelle impulsion d'entrée. Doté d'une entrée unique T, son équation logique est :

$$Q_{n+1} = Q_n \oplus T_n$$

c'est-à-dire qu'un zéro à l'entrée verrouille la sortie dans son état primitif, tandis qu'un 1 provoque l'inversion de l'état des sorties. C'est la raison pour laquelle on l'appelle souvent *diviseur binaire*, et il sera utilisé dans les étages de division.

Fig. 21. — La table d'états de la bascule T.



Son principe est fort simple (fig. 21) : on peut considérer qu'il s'agit d'une variante de JK avec un inverseur supplémentaire à l'entrée unique. De ce fait, ce sont toujours des signaux en opposition qui attaquent les bornes de la bascule proprement dite, l'entrée réelle restant unique.

Dans la figure 22 apparaît le diagramme logique correspondant :

1. Pour $T = 0$, on a forcément $L = M = 0$, donc même avec $K = 1$, on trouve $P = N = 0$. Tout dépend ainsi de l'état antérieur Q_n des sorties :

- Pour $Q_n = 0$ et $\overline{Q_n} = 1$, la porte F fournit $Q_{n+1} = 0$, et pour la même raison, G fournit $\overline{Q_{n+1}} = 1$.
- Pour $Q_n = 1$ et $\overline{Q_n} = 0$, on aboutit à l'inverse. Donc, dans ces deux cas, il n'y a pas eu de changement de l'état des sorties.

2. Pour $T = 1$, on a $K = 0$, donc $P = N = 0$. A nouveau, tout dépend de l'état des sorties :

- Pour $Q_n = 0$ et $\overline{Q_n} = 1$, la porte B se voit appliquer 0 et 1 et fournit donc zéro. Par conséquent, F qui reçoit 3 zéros fournit $Q_{n+1} = 1$. D'autre part, E reçoit 1 et 1 et fournit donc 1, appliqué à G : la sortie Q s'inverse également et passe à $\overline{Q_{n+1}} = 0$.
- Pour $Q_n = 1$ et $\overline{Q_n} = 0$, par symétrie, avec le même raisonnement, on montre que l'état des sorties s'inverse.

Le schéma électrique du flip-flop T peut respecter cette structure (fig. 23). En fait, les MOS permettent, grâce à un fonctionnement dynamique, de la simplifier en mettant les informations en mémoire.

Dans ce dernier schéma, le flip-flop proprement dit est composé des transistors T_1 et T_2 , chargés par T_9 et T_{10} . Quand l'entrée d'horloge CP est positive (niveau logique 0), CP est négatif (niveau 1), rendant conducteurs T_7 et T_8 . Ceux-ci couplent les sorties du flip-flop aux portes T_4 et T_6 , où l'information est stockée dans la capacité de structure gate-substrat.

Si l'on suppose que le drain de T_2 est négatif c'est-à-dire si $Q_n = 1$ et $\overline{Q_n} = 0$, on trouve que le gate de T_6 est négatif et que le gate de T_4 est à la masse. Ces niveaux sont respectivement mémorisés par C_4 et C_6 . Dès lors :

1. Si $\overline{CP} = 1$, les transistors T_7 et T_8 sont conducteurs. Puisque T_4 est bloqué, $A = 1$; mais T_6 étant conducteur, B passe à 0, ce qui produit $Q_{n+1} = 0$ et $\overline{Q_{n+1}} = 1$. Il y a eu inversion. Dans le cas où $Q_n = 0$ et $\overline{Q_n} = 1$, l'inversion symétrique se produirait.

2. Si $CP = 1$, c'est-à-dire si $\overline{CP} = 0$, les transistors T_7 et T_8 sont bloqués, ainsi d'ailleurs que T_3 et T_5 . Par conséquent, A et B non plus que Q et \overline{Q} ne peuvent changer d'état.

On a appelé ici CP l'entrée T; en effet, comme ce flip-flop sert de diviseur par 2, on lui applique à l'entrée des impulsions d'horloge (CP vient de « *clock pulse* »).

6. LE FLIP-FLOP D

Le D à verrouillage

La bascule D est un élément de recopie : la sortie suit l'entrée, après un *décalage* (D) dû au temps de propagation. Par conséquent, sa formule logique est $Q = D$ et $\overline{Q} = \overline{D}$, et sa structure est très simple (fig. 24); en se référant aux tables des AND et NOR rappelées pour mémoire dans la même figure, et en tenant compte de l'entrée d'horloge \overline{H} , on voit que :

- Si $D = 0$, on a $K = 1$ et, lorsque l'horloge est haute ($H = 1$), L passe à 1 et M = 0. La sortie du NOR noté F, pour une entrée à 1, est forcément $Q = 0$ et, par conséquent, E qui se voit appliquer deux zéros donne $\overline{Q} = 1$.

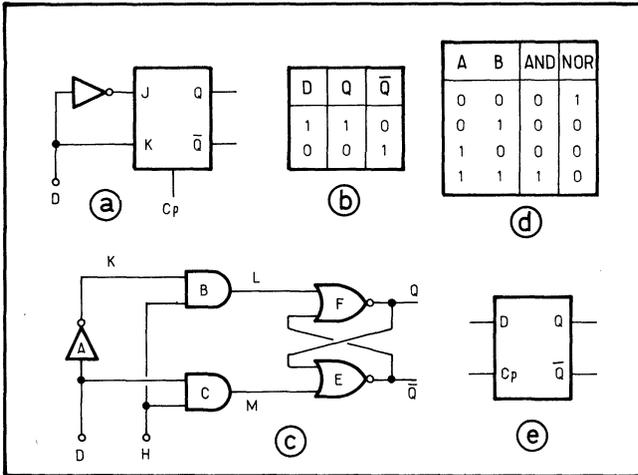
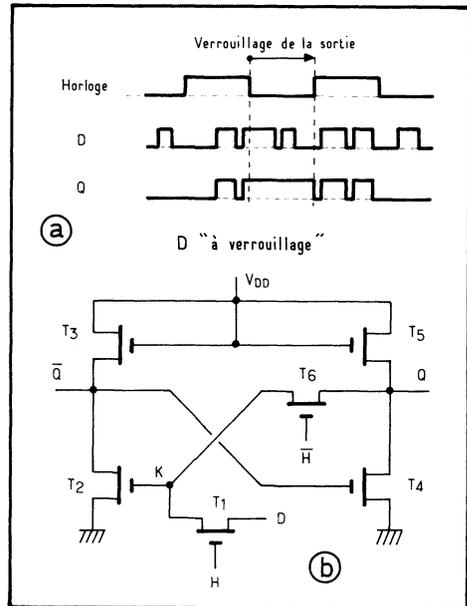


Fig. 24. — Le flip-flop D: principe (a), table d'états (b), réalisation (c);

Fig. 25. — Le flip-flop à verrouillage.



— Si $D = 1$ et $K = 0$, on obtient, lorsque $H = 1$, un 0 en L et un 1 en M. Par conséquent, E donne forcément $\bar{Q} = 0$, et F fournit en conséquence $Q = 1$.

Une telle bascule D est dite à verrouillage (bistable latch), car tant que l'horloge est haute, la sortie suit l'entrée; mais le front arrière de l'impulsion d'horloge verrouille les sorties sur la dernière situation acquise (fig. 25). Son schéma électrique est relativement simple (même figure), car les portes AND sont d'unique transistors; ainsi, T₁ est un AND avec D et H, etc.

Le D à commutation sur les fronts

Il existe cependant une variante de ce D, où le verrouillage est obtenu sur le front avant de l'impulsion d'horloge, ce qui mène aux signaux de la figure 26. Un tel flip-flop est appelé D à commutation sur les fronts, car c'est la tension croissante d'horloge qui provoque, à un certain niveau, le verrouillage des sorties.

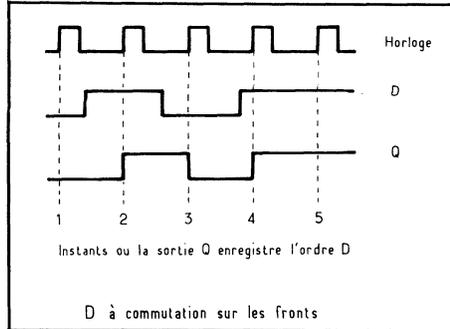


Fig. 26. — Chronogramme du D à commutation sur les fronts.

- Pour $H = 0$, K reste au niveau acquis; T_6 étant conducteur, le flip-flop maintient son état antérieur.

- Pour $H = 1$, le transistor T_6 se bloque et coupe la liaison \overline{Q} à K ; par contre, T_1 devient conducteur et l'on retrouve deux cas :

1. Pour $D = 0$, on obtient $K = 0$ donc T_2 se bloque ce qui fournit $\overline{Q} = 1$. Via T_4 débloqué, Q passe par conséquent à 0.

2. Pour $D = 1$, il est évident qu'on obtient $K = 1$, puis $\overline{Q} = 0$ et, par conséquent, $Q = 1$.

En fait, le D à commutation sur les fronts est un flip-flop du type maître-esclave.

7. LA BASCULE RST

La bascule RST résulte de la combinaison d'un RS et d'un T. Par conséquent, sa table des états (fig. 27) doit révéler :

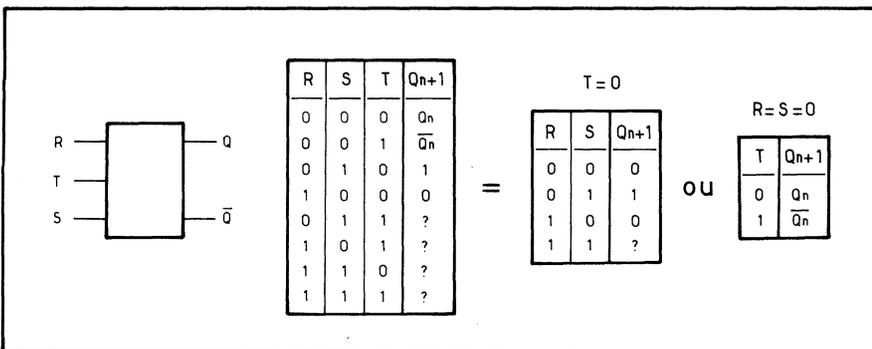


Fig. 27. — La bascule RST : ses tables d'états.

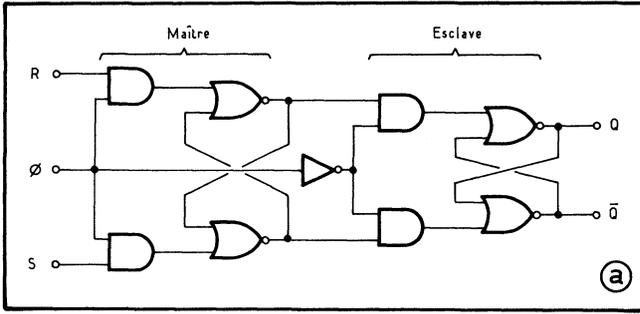
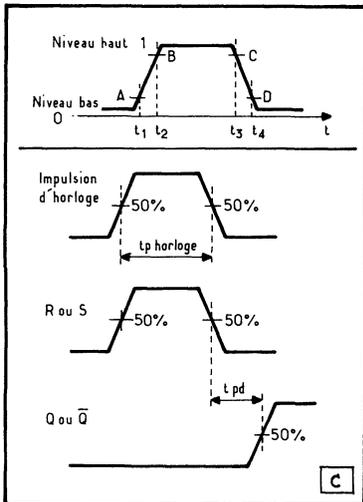
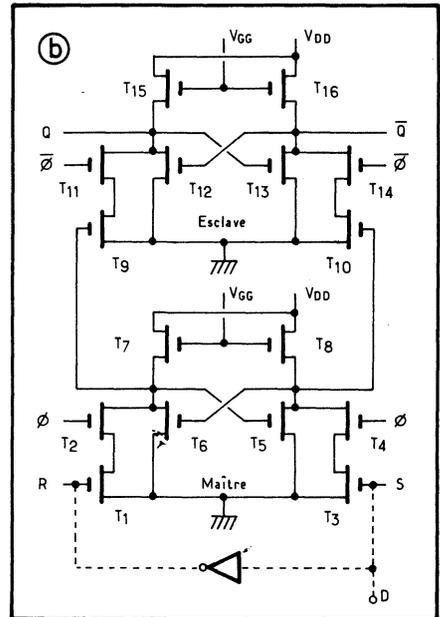


Fig. 28. — Le maître-esclave : schémas logique (a) et électrique (b). En c, séquence de fonctionnement : en t_1 , la bascule maître est isolée de la bascule esclave; en t_2 , le maître enregistre l'information; en t_3 , les entrées J et K sont fermées; en t_4 , l'esclave enregistre l'information transmise par le maître. En (b), les conditions de synchronisme entre les entrées J et K, et le signal d'horloge d'une telle bascule.



- un fonctionnement en RS lorsque $T = 0$;
- un fonctionnement en T, c'est-à-dire que chaque 1 appliqué à l'entrée T provoque un basculement à la sortie, lorsqu'on fait $R = S = 0$.

On ne doit pas confondre le RST avec le RS synchrone.

8. LES BASCULES MAITRE-ESCLAVE

Les *maître-esclave* sont constitués, en réalité, par deux bascules en série, l'une commandant l'autre, tels les deux RS de la figure 28 a. Ces RS synchrones sont exactement semblables à ceux qui ont été présentés antérieurement, mais ils sont attaqués par des phases d'horloge inversées.

On voit ainsi que lorsque $\Phi = 1$, le RS *maître* enregistre les ordres présents à l'entrée, mais ne les transmet à l'*esclave* que lorsque Φ retourne à zéro. Ce n'est qu'à ce moment que les décisions apparaîtront aux sorties.

La séquence de fonctionnement est représentée sur la figure 28, en prenant pour référence l'impulsion d'horloge. A l'instant t_1 , la bascule maître est isolée de l'esclave; à l'instant t_2 , la bascule maître enregistre l'information aux entrées R et S. A l'instant t_3 , les entrées R et S sont fermées pour éviter un nouveau changement de la bascule maître; à l'instant t_4 , les bascules maître et esclave sont réunies, l'information est alors transférée du maître à l'esclave et apparaît aux sorties Q et \bar{Q} .

Cette bascule présente certaines caractéristiques intéressantes. En effet, un de ses avantages réside dans le fait qu'elle n'est pas sensible aux temps de montée et de descente des signaux; cela résulte du fait que cette bascule commute sur des niveaux de tension avec des seuils séparant l'état 1 et l'état 0.

Enfin, une autre particularité est que les informations, sur les entrées R et S, doivent être présentes pendant toute la durée de l'impulsion d'horloge; mais il n'est pas nécessaire d'avoir un préconditionnement ou un temps de maintien.

La structure d'un RS maître-esclave est donnée figure 28 b, en développant le synoptique précédant. En pointillé, on a placé un inverseur qui le transformerait en D maître-esclave; on n'a cependant pas représenté l'inverseur d'horloge, les entrées d'horloge de même phase étant bien entendu réunies.

En recourant à des flip-flops JK ou T, on obtiendrait de la même façon des JK maître-esclave ou des T maître-esclave.

9. RÔLE DES BORNES « PRESET » ET « CLEAR »

Diverses bascules sont dotées d'entrées annexes, dont les bornes *preset* et *clear*, appelées selon les circonstances *set direct* ou *clear direct*, *reset* ou *preclear*, ou en français *RAZ* pour remise à zéro, et *RA 1* pour remise à 1 : elles servent à imposer l'état 0 ou 1 à la sortie Q de façon prioritaire, même en l'absence de signal d'horloge, donc de façon asynchrone.

10. LES FLIP-FLOPS EN CMOS

Avec les MOS complémentaires et en appliquant les mêmes principes, on aboutit à des structures de flip-flops dont les schémas suivants donnent quelques exemples.

Les RS en CMOS

On a vu, au début de ce chapitre, la structure du RS à portes NOR, en CMOS. Avec des portes NAND, on aboutit aux schémas de la figure 29; à nouveau, il s'agit là d'un montage réel correspondant au circuit type *CD 4044 A*, doté d'une sortie 3 états et de réseaux de protection des entrées.

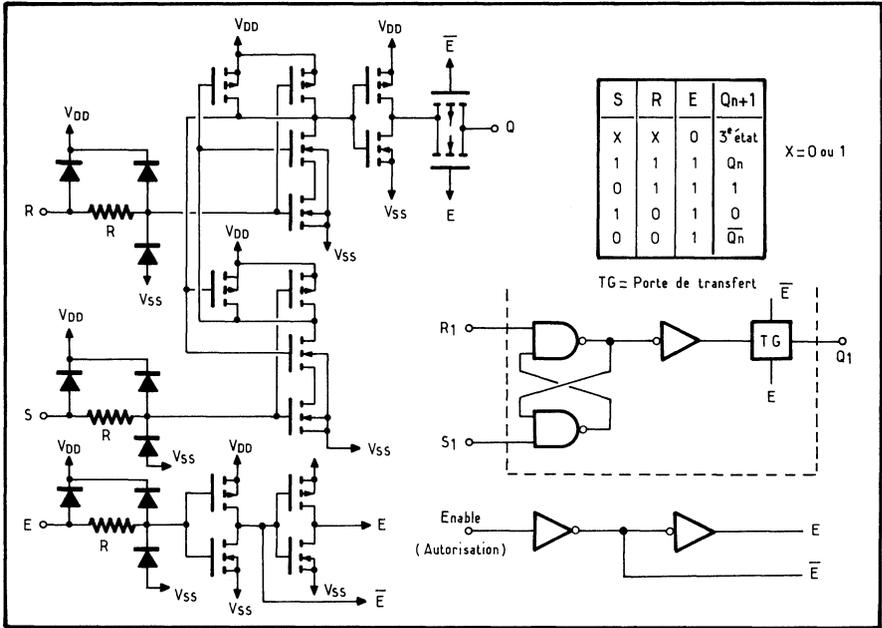


Fig. 29. — Un RS à portes NAND en CMOS.

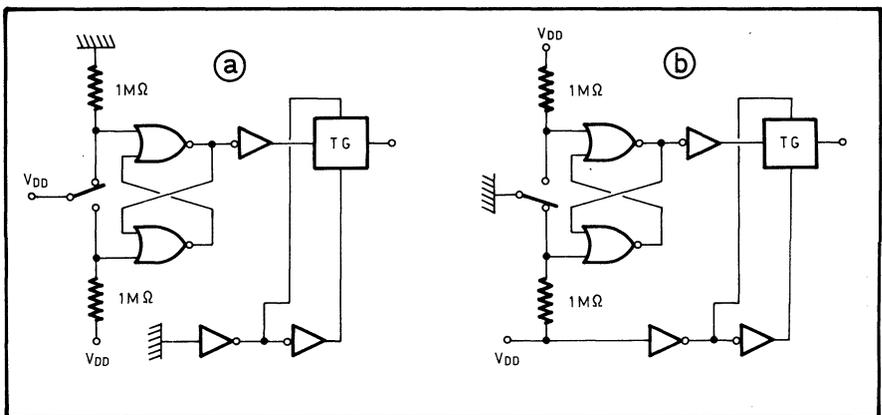


Fig. 30. — Utilisation du RS en commutateur sans rebondissement.

Les 4043 A (portes NOR) et 4044 A acceptant des V_{DD} de 3 à 15 V, consomment quelque 5 nA au repos et pilotent environ 0,4 mA en sortie. Leurs temps typiques sont de $t_{PHL} = 175$ ns, $t_{PLH} = 75$ ns, $t_{THL} = 100$ ns et $t_{TLH} = 50$ ns, pour une capacité d'entrée de 5 pF. Elles peuvent ainsi fonctionner jusqu'à 10 MHz (sous 15 V). La figure 30 montre leur montage classique en commutateur sans rebondissement.

Les flip-flops D

En fait, le D a davantage la structure d'étages de registre à décalage que celle de flip-flops classiques à rétro-action croisée (fig. 31), et ne comporte que des inverseurs et des portes de transfert notées PT, ou encore TG, de *transfer gate*.

Fig. 31. — Un demi flip-flop D, tel que proposé par RCA en C/MOS; les notations PT indiquent les portes de transfert (en a). Le flip-flop complet, maître-esclave, réunit deux de tels ensembles; les horloges sont inversées dans le maître et dans l'esclave (en b).

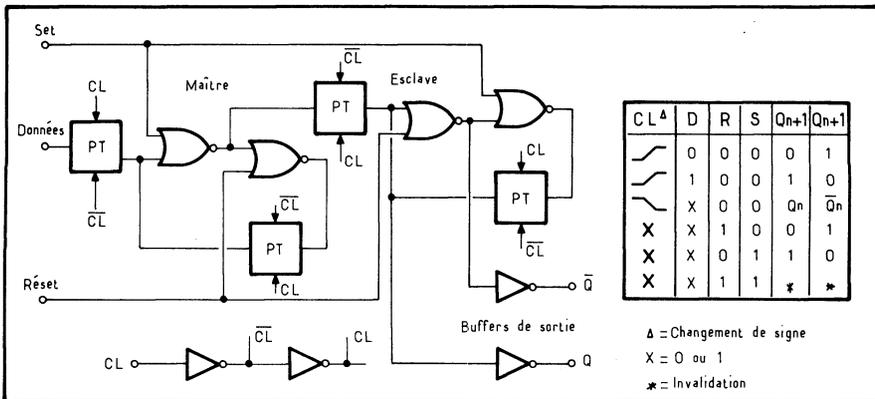
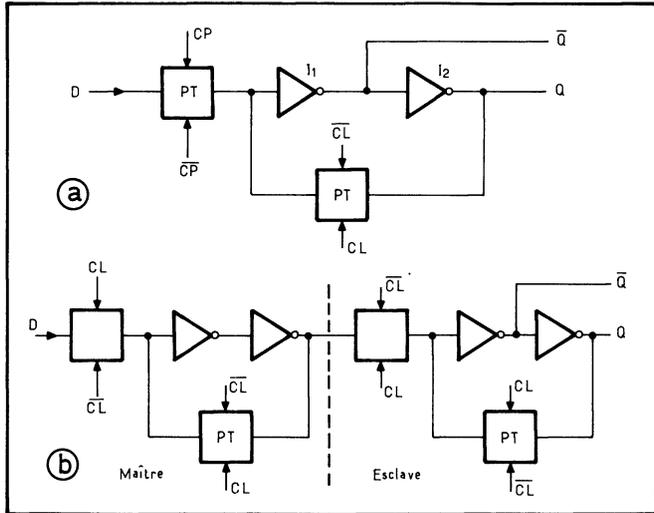
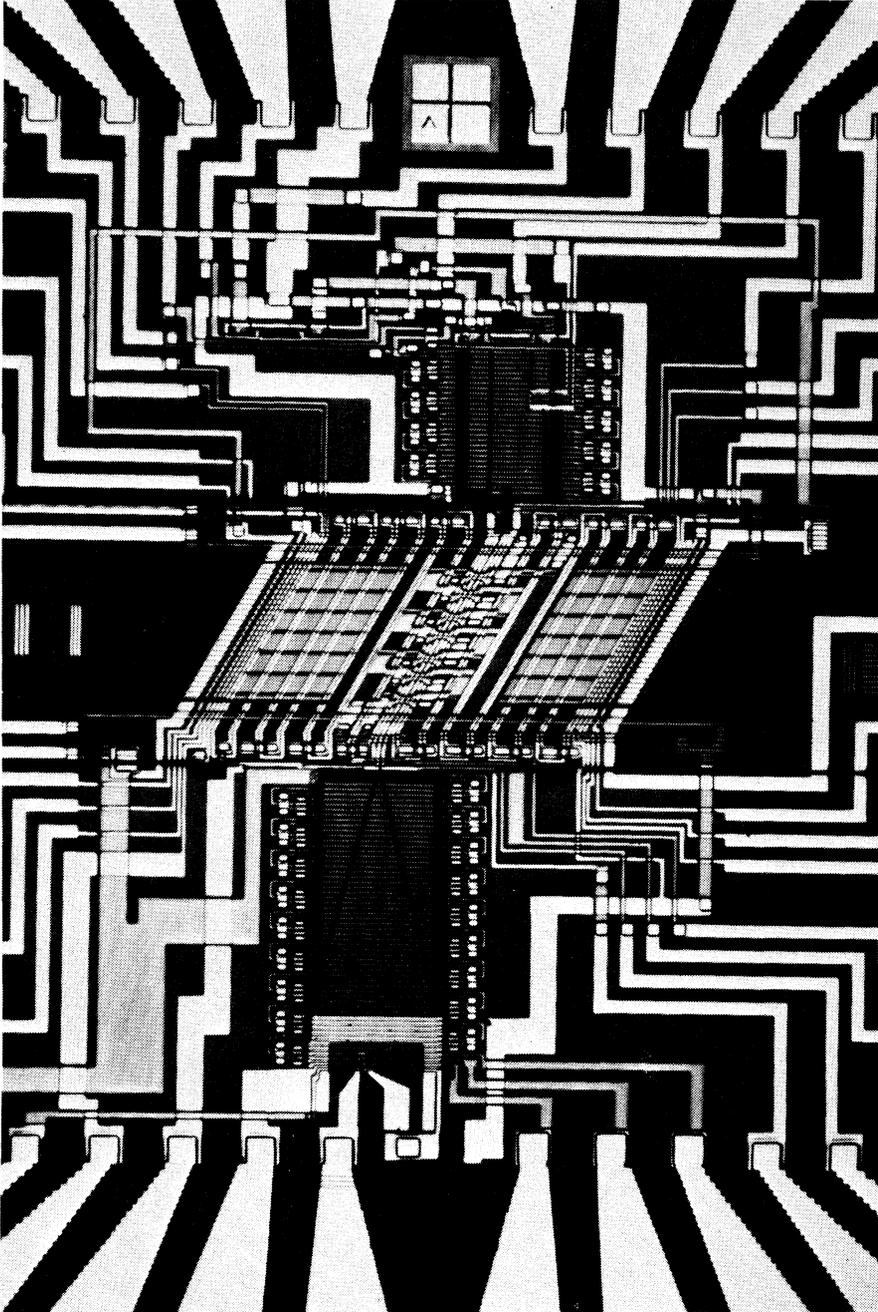


Fig. 32. — Comment s'effectue l'adjonction des commandes *remise à zéro* et *remise à un* (set et reset) sur le FF maître-esclave D de la figure précédente. L'impulsion d'horloge CL attaque un double inverseur en série.

Leur fonctionnement est évident, en D simple ou en maître-esclave. Il ne faudra recourir à des NOR que si l'on veut adjoindre au D des commandes de *remise à zéro* (RAZ, ou *reset*), et de *remise à un* (RA 1, ou *set*), comme le montre la figure 32 pour un D maître-esclave (type *CD 4013*). L'horloge y est notée CL (de « clock pulse »).

On voit que, dans les deux cas, la sortie du maître suit l'entrée D tant que l'horloge est haute, et que la sortie est verrouillée, et transmise à l'esclave, dès que l'horloge retourne au niveau bas. Les temps typiques de tels éléments varient autour de 100 ns, montée et descente, transition ou propagation. Il s'agit là, en fait, du véritable D à commutation sur les fronts.



Réalisée selon des techniques relevant de l'intégration, les mémoires à bulles magnétiques ont été commercialisées pour la première fois en 1977. On voit sur ce document IBM la topologie d'une telle mémoire, à densité de stockage de un million de bits par centimètre carré (pour les bulles seules).

CHAPITRE XII

LES REGISTRES

Les registres constituent un chapitre important des fonctions MOS. En effet, outre leurs applications à la conversion et au comptage, ils débouchent directement sur les mémoires. Ils se partagent en registres statiques ou dynamiques.

1. PRINCIPE DES REGISTRES A DÉCALAGE

Un registre à décalage se compose d'une série de cellules élémentaires (contenant une information, 0 ou 1), mises successivement en communication sur ordre de l'horloge. Les informations se déplacent alors d'un cran, d'une cellule à la suivante, à chaque impulsion d'horloge. Chaque cellule chargée de stocker un bit peut être un flip-flop ou une porte dynamique.

Supposons que ces cellules soient en parallèle; dans la figure 1 *a*, elles sont repérées A, B, C et D. Les interrupteurs représentent, en fait, des MOS rendus conducteurs à l'application de l'impulsion d'horloge. Si l'on ne recourait qu'à une seule phase d'horloge, la ligne serait court-circuitée et toutes les informations « glisseraient » jusqu'à la sortie; par contre, deux phases d'horloge en opposition mettront les cellules en contact deux à deux, un cycle d'horloge complet exigeant alors deux temps.

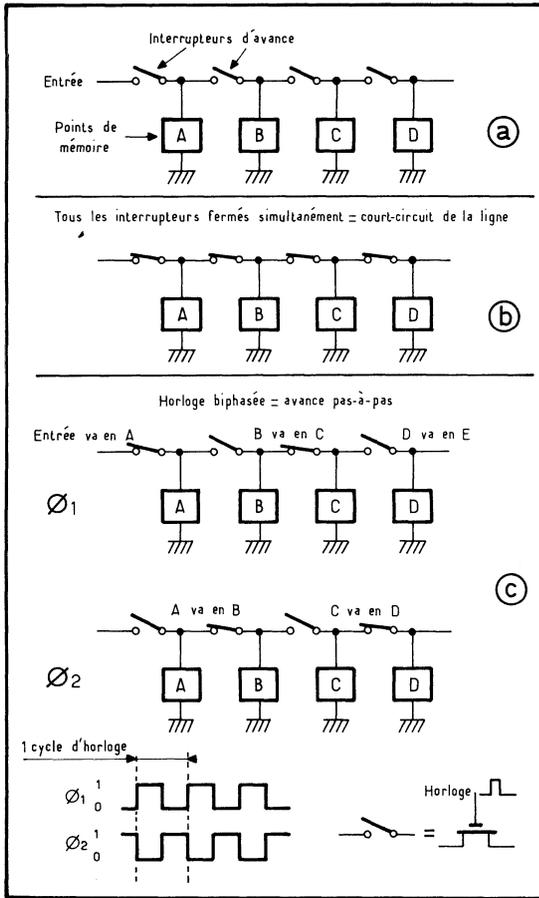


Fig. 1. — Le principe des registres à décalage.

Telle est bien la conception de base des registres à décalage; elle pourra être modifiée dans deux directions :

1. On utilisera une seule phase d'horloge lorsque les cellules seront des bascules statiques où l'information n'apparaît aux sorties que sur le front descendant de l'horloge. C'est le cas du maître-esclave en particulier.

2. On utilisera 4 phases d'horloge pour des cellules dynamiques inverseuses du type biphasé.

2. LES TYPES DE REGISTRES

Un registre se caractérise par sa longueur, en nombre de cellules, c'est-à-dire en nombre de bits. Ainsi, le registre de la figure 2 est un élément 4 bits.

Fig. 2. — Principe d'un registre à décalage statique, recourant à des flip-flops.

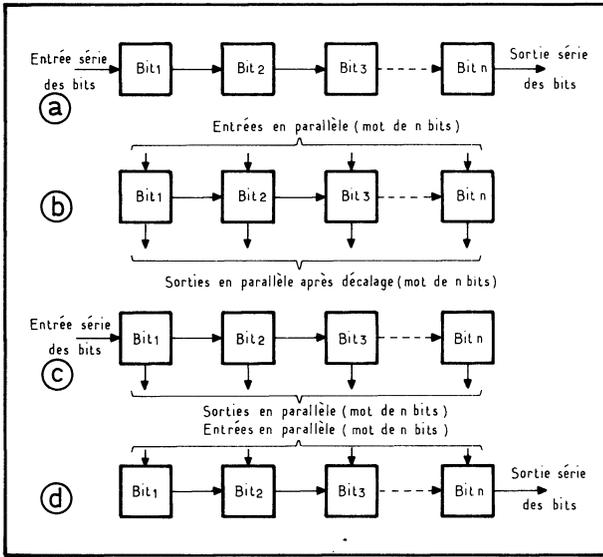
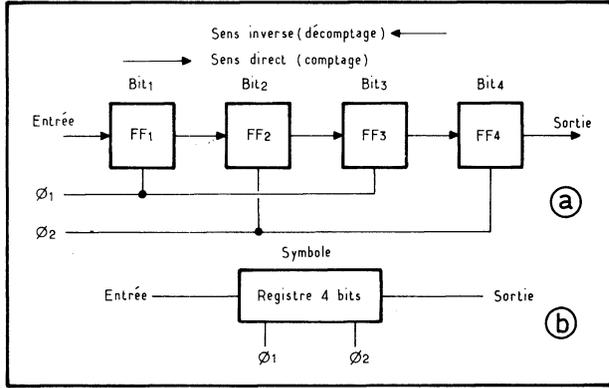


Fig. 3. — Les 4 possibilités d'entrées-sorties.

Par convention, on suppose que l'information circule de la gauche (*entrée*) vers la droite (*sortie*) ; c'est le sens du *comptage*. Par opposition, le sens inverse ou *rétrograde* s'appellera *décomptage*. Il existe des registres où le sens de comptage est commandé à l'aide d'un ordre spécial; ce sont des registres dits *réversibles*.

L'introduction comme l'extraction des informations — des bits — ne se fait pas obligatoirement en série. Les quatre possibilités sont (fig. 3) :

- *Entrée série, sortie série*

Dans le registre à décalage à entrée série et sortie série, la première impulsion de commande introduit la première information dans la bascule 1. La seconde impulsion de décalage déclenche le transfert de cette information de 1 en 2

et introduit la seconde information en 1. La troisième impulsion de décalage poursuit ce cycle, ainsi que les suivantes et cela, jusqu'à ce que ces informations parviennent à la sortie (en série).

• *Entrée parallèle, sortie parallèle*

L'introduction et l'extraction des données se font, dans ce cas, en mode parallèle et le plus souvent par mot (de n bits) complet. Une seule impulsion d'horloge est donc suffisante pour provoquer l'entrée du mot dans le registre. L'extraction peut se faire sans, ou après décalage.

• *Entrée série, sortie parallèle*

Il s'agit là d'une nouvelle variante, où les données sont introduites en série et lues en parallèle.

• *Entrée parallèle, sortie série*

Là, c'est la situation inverse qui prévaut. En effet, les mots binaires sont introduits en parallèle et parviennent en série à la sortie grâce aux impulsions d'horloge qui commandent leur décalage pas à pas.

Enfin, on trouvera des registres :

- à longueur variable;
- rebouclés,

sur lesquels on reviendra ultérieurement. Ils seront réalisés avec des cellules statiques ou dynamiques dont la figure 4 illustre l'action, avec d'ailleurs la fonction « effacement » : elle consiste à appliquer une unique phase d'horloge simultanément à toutes les cellules. On déduit de ce dessin que :

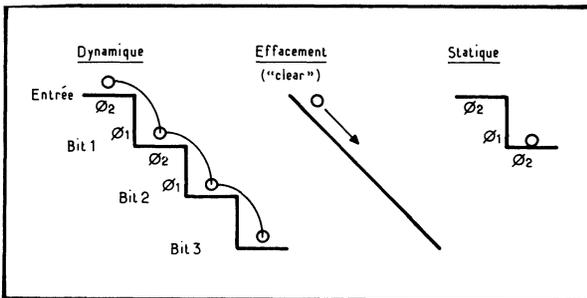


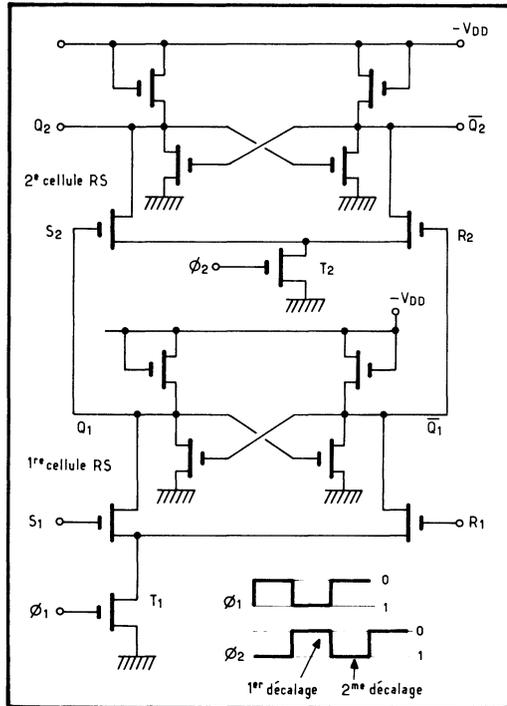
Fig. 4. — Illustration du fonctionnement des registres.

- le registre dynamique n'accepte qu'une fréquence minimale donnée de fonctionnement. En deçà de cette fréquence, la durée d'une période excède les possibilités de stockage de la capacité intrinsèque des MOS, qui se décharge et par conséquent perd l'information. Cette fréquence minimale est de l'ordre de 10 kHz. La fréquence maximale de travail va, elle, jusque quelque 10 MHz;
- le registre statique peut travailler à la fréquence nulle car sa cellule, un flip-flop, conserve l'information. Par contre, sa structure étant plus complexe, il fonctionne moins vite. Ainsi, sa fréquence maximale sera de l'ordre de 3 MHz.

3. LES REGISTRES STATIQUES BIPHASES

Les registres statiques biphasés sont réalisés par l'association d'autant de paires de cellules statiques qu'on désire de bits. Ainsi, la figure 5 montre (avec une disposition verticale pour plus de clarté) deux bascules RS connectées pour former un registre de 1 bit; elles sont, par conséquent, attaquées par deux phases d'horloge. On pourrait d'ailleurs assimiler un tel montage à un classique maître-esclave.

Fig. 5. — Deux RS en série constituent un registre de 1 bit de délai, puisqu'il faut deux phases d'horloge.



Il faut donc deux RS pour réaliser une cellule dite de 1 bit de délai; il est évident, en effet, qu'une bascule suivante doit avoir été libérée au préalable de l'information qu'elle contenait pour recevoir ensuite l'information venant de la cellule qui la précède. Donc, un cycle complet d'horloge de deux phases est nécessaire pour ce délai de 1 bit. On en déduit qu'un maître-esclave correspond également à un délai de 1 bit.

Tout autre type de bascule statique peut s'appliquer à un tel montage. En pratique, d'ailleurs, on trouvera généralement les structures les plus dépouillées, telles que des inverseurs rétro-couplés.

4. LES REGISTRES DYNAMIQUES BIPHASES A INVERSEURS PROPORTIONNELS

Le recours à de simples inverseurs proportionnels est illustré par le schéma de principe d'un registre dynamique biphasé (fig. 6). Dans ce montage, la consom-

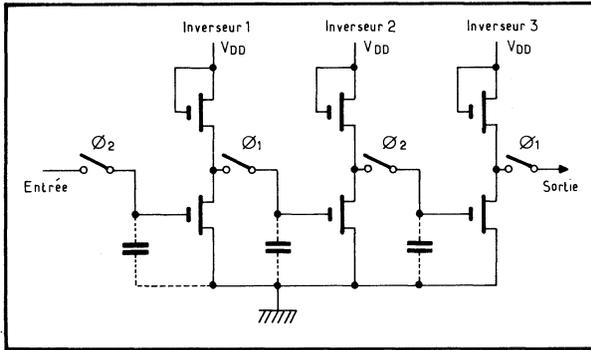
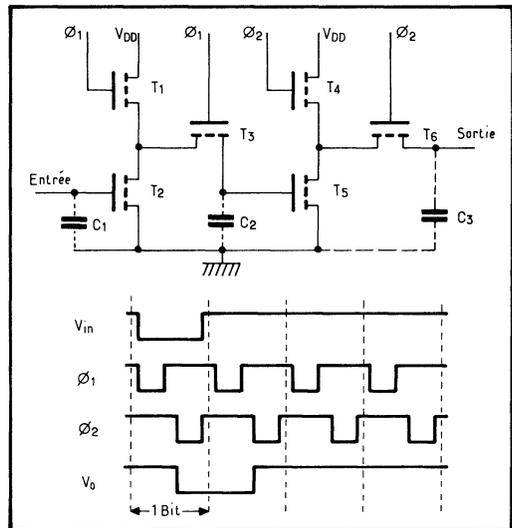


Fig. 6. — Principe du registre dynamique biphasé.

Fig. 7. — Schéma électrique du registre dynamique biphasé.



mation est relativement importante lorsque le MOS inverseur est conducteur, car la charge reste en service en permanence. Il mène au schéma électrique de la figure 7, pour un délai de 1 bit.

Une autre méthode procurant cependant un faible isolement inter-étages mais réduisant la consommation en bloquant la charge un temps sur deux est indiquée figure 8; les phases d'horloge sont appliquées aux inverseurs. Le fonctionnement en est le suivant.

En t_1 , un niveau 1 est appliqué à l'entrée et charge la capacité de structure C_A . En t_2 , Φ_1 débloque Q_2 ; de ce fait Q_1 et Q_2 conduisent et le potentiel, en B, est au niveau 0.

En t_3 , Φ_2 débloque Q_4 , mais puisque le point B est à la masse, Q_3 ne conduit pas; la capacité C_C se charge donc à $-V_{DD}$ (niveau 1). La sortie a alors suivi l'entrée, le temps du transit correspond à l'espace t_1-t_3 .

Fig. 8. — Structure de principe d'un registre à décalage dynamique (ϕ); signaux typiques recueillis en divers points du montage (b).

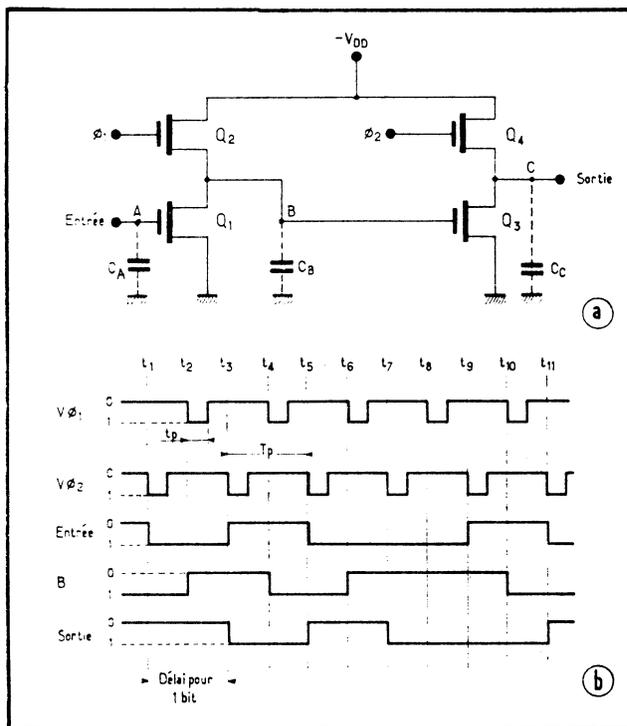
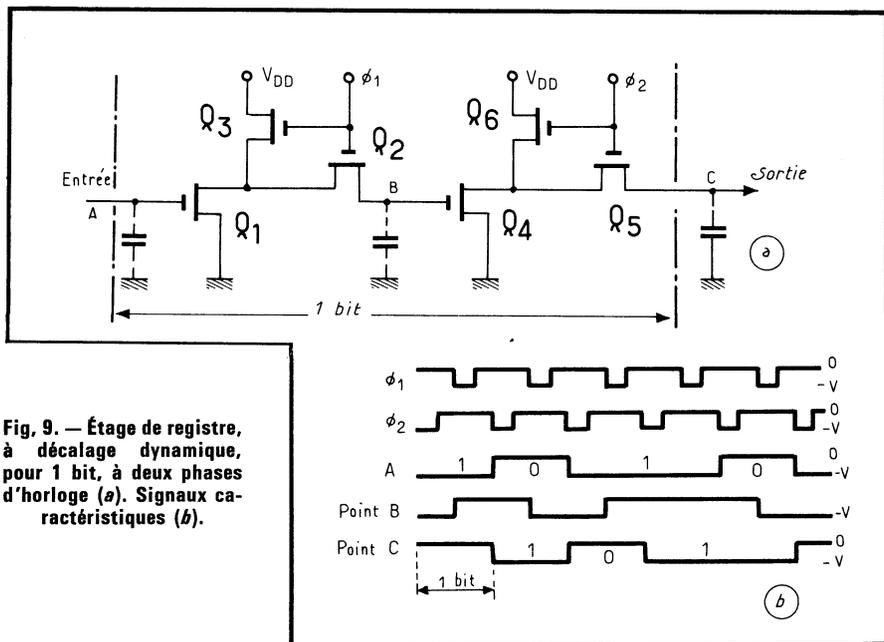


Fig. 9. — Étage de registre, à décalage dynamique, pour 1 bit, à deux phases d'horloge (a). Signaux caractéristiques (b).



L'analogie est donc grande entre cet étage et un flip-flop de type D. Cependant, l'isolement inter-étages peut être amélioré, comme le montre le schéma de la figure suivante (fig. 9) qui reprend, avec un dessin différent, le montage déjà donné en 7. On retrouve ici les types d'inverseurs étudiés antérieurement. On voit que les transistors Q_2 et Q_5 fonctionnent en direct pour charger les capacités de structure via Q_3 et Q_6 , en inverse pour les décharger via Q_1 et Q_4 . Les inconvénients sont encore multiples :

- puisque deux MOS en série sont toujours conducteurs, le niveau du zéro se dégrade;
- la rapidité est moindre;
- un courant important traverse chaque inverseur lorsqu'il est conducteur.

Les bascules RS semi-dynamiques, ou D, telles qu'étudiées dans le chapitre précédent, constituent elles aussi des étages de registres à décalage. Les figures 10

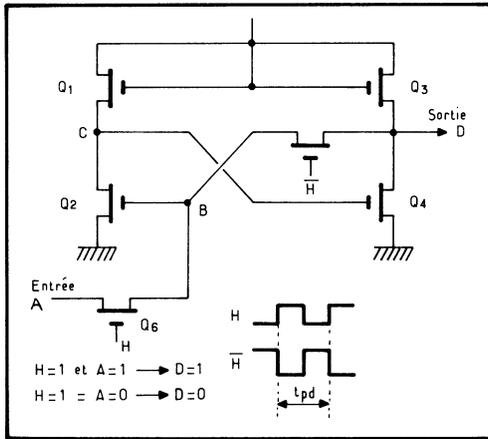


Fig. 10. — Une bascule D en cellule de registre.

et 11 en donnent deux exemples; le fonctionnement de ce dernier montage est le suivant :

- en t_1 , $\Phi_2 = 0$; par conséquent, Q_3 et Q_4 sont bloqués et C_1 et C_2 ne changent pas d'état. La sortie ne change pas;
- en t_2 , Φ_1 passe à 1 et Q_5 se débloque; l'information présente en A est transmise à B. Si A = 1, Q_1 devient conducteur et C passe à la masse;
- en t_3 , $\Phi_1 = 0$; Q_5 se bloque et C_1 maintient Q_1 dans l'état précédent;
- en t_4 , $\Phi_2 = 1$ et met en service Q_3 et Q_4 qui verrouillent la bascule, devenue alors statique.

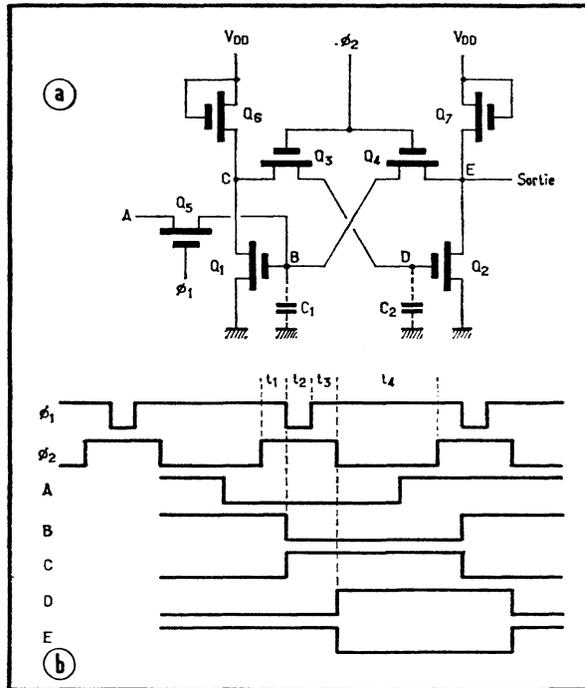
5. REGISTRE DYNAMIQUE PROPORTIONNEL TRIPHASE

Le fonctionnement de principe du registre statique à 3 phases d'horloge est identique au précédent. Ces 3 phases sont, sur le schéma de la figure 12, appelées Φ_1 , Φ_2 et Φ_3 .

Supposons, au préalable, que le flip-flop a emmagasiné un 0, donc que T_5 est débloquenté et que la sortie est à la masse (approximativement), donc au niveau 0 correspondant à une tension plus positive que -1 V. Le gate de T_5 est alors au niveau 1, plus négatif que -10 V.

En t_1 , Φ_1 passe à 1 et T_1 est débloquenté. Simultanément, Φ_3 passe à 0, bloquant T_2 et isolant le point A de la sortie du flip-flop. L'entrée étant au niveau 1, A passe à 1 et T_4 se débloquenté; de ce fait, le point B passe à 0. Comme Φ_2 est également à 0, T_3 est bloqué et le point B est isolé du point C.

Fig. 11. — Un RS semi-dynamique en registre.



Cependant, T_5 reste passant tant que la capacité intrinsèque du MOS n'est pas déchargée et il maintient la sortie du flip-flop à 0. Précisons que l'impulsion d'horloge, en t_1 , ne doit pas excéder $20 \mu s$, afin que cette capacité n'ait pas le temps de se décharger sensiblement.

A l'instant t_2 , Φ_1 passe à 0 tandis que Φ_2 , suivi par Φ_3 , passe à 1. Dès que $\Phi_1 = 0$, le MOS T_1 se bloque et l'entrée est isolée de la bascule; cependant, T_4 reste conducteur en raison de la charge antérieure de sa capacité intrinsèque.

Donc, à l'instant où Φ_1 passe à 0 et Φ_2 à 1, le MOS T_3 se débloquenté. Le point B étant bas, la capacité de structure, en B, se décharge via T_3 et le MOS T_5 se bloque. Puis, l'impulsion d'horloge retardée Φ_3 passe à 1, débloquenté T_2 et connectant le point A à $-V_{DD}$, via R_2 ; la capacité de structure de T_4 se charge et ce MOS est rendu conducteur.

L'impulsion retardée Φ_3 est donc indispensable dans ce cas, car sinon, T_2 devient conducteur en même temps que T_3 : nul ne peut alors prédire si T_5 se bloquera avant que T_4 ne devienne conducteur, ou si c'est l'inverse qui se produira. A la limite, le montage se mettrait à osciller.

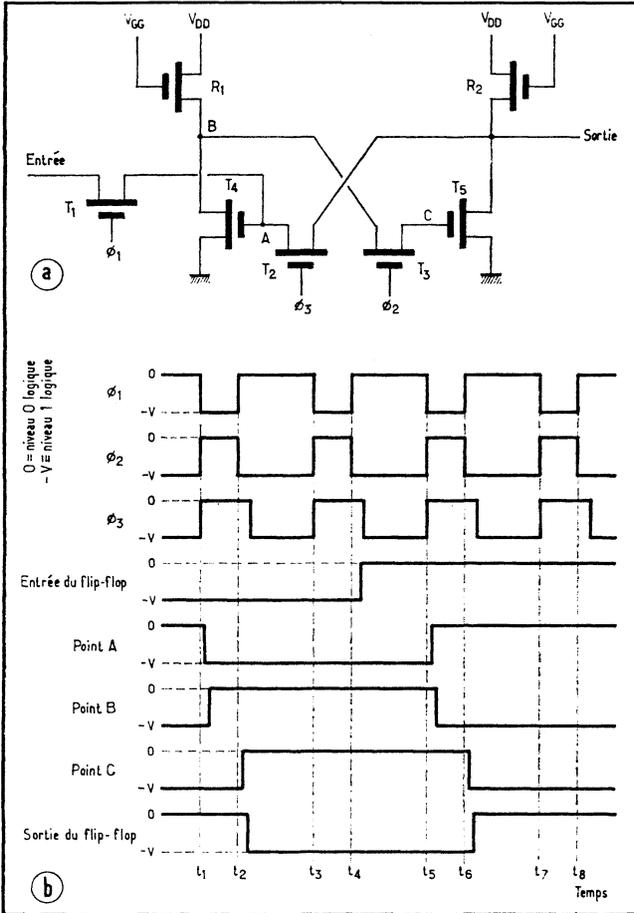


Fig. 12. — Registre à décalage statique à horloge triphase (a); quelques signaux typiques (b).

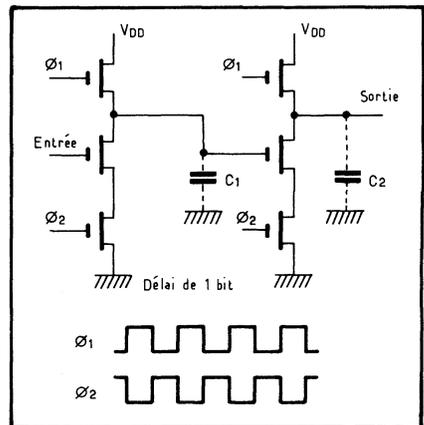


Fig. 13. — Principe du registre dynamique non proportionnel à précharge.

A l'instant t_3 , Φ_1 repasse à 1 tandis que Φ_2 et Φ_3 repassent à 0; l'entrée conservant un niveau 1 (comme le montre le diagramme), T_4 reste conducteur et le point B est maintenu au niveau bas. A partir de l'instant t_4 , les signaux d'horloge changent à nouveau d'état mais le flip-flop reste dans une situation inchangée (T_4 conducteur et T_5 bloqué).

Lors de l'impulsion d'horloge suivante, en t_5 , l'entrée est attaquée par un 0 logique. La capacité de structure au point A est alors déchargée via T_1 et T_4 se bloque. Le point B passe au niveau haut et, à l'instant t_6 , quand T_3 devient conducteur, le point C passe au niveau haut : T_5 devient passant et la sortie est 0.

Lorsque Φ_3 passe à 1, T_2 est rendu conducteur et T_4 est maintenu bloqué. Enfin, la dernière impulsion d'horloge de la séquence, à l'instant t_7 , débloque T_1 . L'entrée étant à 0, le MOS T_4 reste bloqué et B est maintenu au niveau haut. En t_8 , rien ne vient plus modifier cet état de choses et la sortie reste à 0.

6. REGISTRES DYNAMIQUES NON PROPORTIONNELS A PRÉCHARGE

Ainsi qu'on l'a vu, l'inverseur dynamique non proportionnel n'offre aucun chemin continu, à aucun moment, au courant entre les lignes d'alimentation.

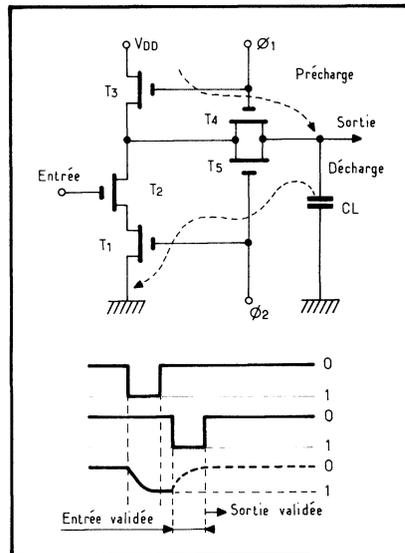


Fig. 14. — Registre à précharge avec porte de transfert.

Son utilisation se traduira par une économie en courant. Le schéma de principe d'un registre non proportionnel est donné figure 13 :

- Φ_1 charge les capacités de structure $C_1, C_2...$ On dit que cette phase *précharge* les circuits;
- Φ_2 , si l'ordre d'entrée est 0, maintient en l'état de niveau de charge. Si l'ordre d'entrée est 1, cette phase décharge les capacités de structure.

L'isolement inter-étages est cependant faible. Si on veut l'améliorer, on recourt au montage suivant (fig. 14), où l'on tient compte également de l'existence de capacités drain-substrat (dont on accroît d'ailleurs la valeur de fabrication. Ainsi :

- Φ_2 précharge C_2 et charge C_1 au niveau d'entrée;
- Φ_1 décharge C_2 si l'entrée = 1; si l'entrée = 0, la charge de C_2 est transférée à C_3 .

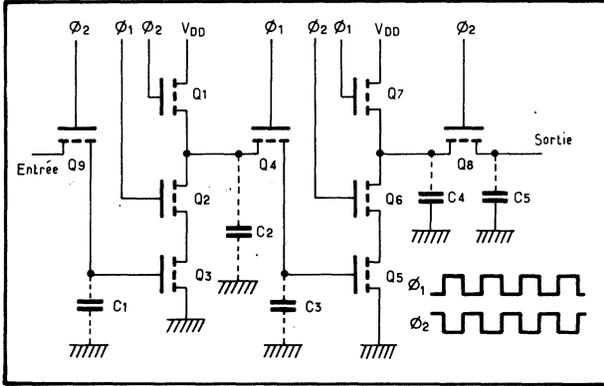
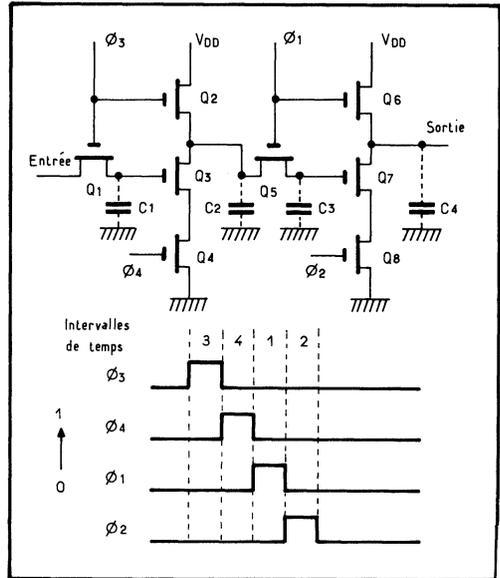


Fig. 15. — Registre à précharge avec double porte de transfert.

Fig. 16. — Registre quadriphase à précharge.



On remarquera une inversion systématique des phases appliquées aux deux étages. On pourrait utiliser une double porte de transfert (fig. 15), ou encore recourir à 4 phases d'horloge (fig. 16) :

- Φ_3 débloque Q_1 et charge C_1 au niveau d'entrée, d'une part, et Q_2 qui précharge C_2 à V_{DD} .

Fig. 17. — Registre quadriphase avec séparation des fonctions de charge et décharge.

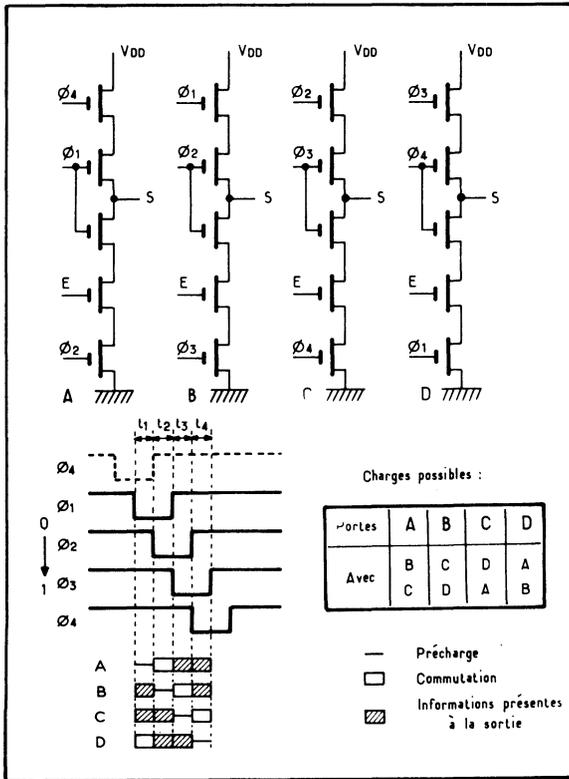
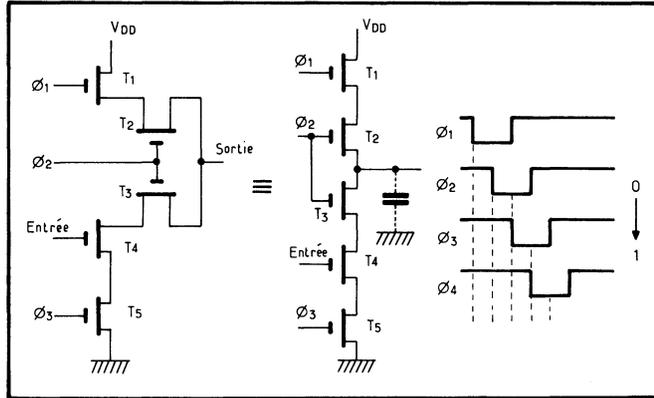


Fig. 18. — Autre montage quadriphase commandé par des impulsions d'horloge d'égale durée.

- Φ_4 débloque Q_4 qui décharge C_2 si Q_3 conduit (C_1 étant à 1);
- Φ_1 transmet le niveau de C_2 à C_3 ;
- Φ_2 agit comme Φ_4 .

On aboutit à un fonctionnement nettement segmenté. A noter qu'il suffirait de reboucler l'entrée sur la sortie pour obtenir une bascule statique.

Au prix d'un MOS supplémentaire, on peut modifier cette porte pour mieux séparer les fonctions de charge et de décharge (fig. 17); deux MOS apparaissent pour la porte de transfert; la phase Φ_4 est utilisée avec l'inverseur suivant. Ici, les impulsions d'horloge non seulement peuvent se chevaucher mais, pour une fréquence identique, occupent une longueur de temps double.

La succession de telles portes dans un registre obéit à des règles strictes pour adapter les diverses phases du fonctionnement. C'est ce que montre la figure 18 où, si l'on considère la structure A :

- en t_1 , $\Phi_4 = \Phi_1 = 1$ préchargent la sortie;
- en t_2 , $\Phi_1 = \Phi_2 = 1$ déchargent la sortie si l'entrée est à 1;
- en t_3 et t_4 , les informations sont disponibles à la sortie.

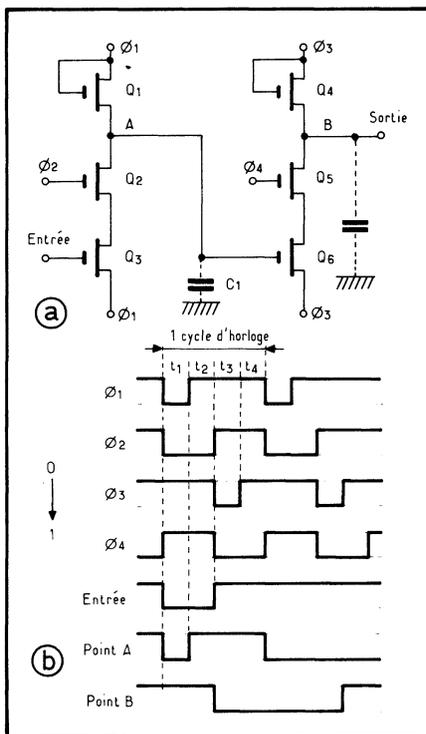


Fig. 19. — Variante du quadriphase à phases décalées.

De ce fait, la porte A ne peut être suivie que par B ou C, puisqu'alors la phase de commutation correspond à la présence des informations à la sortie de A (temps t_3 et t_4).

Une variante est dessinée figure 19; si l'on considère la porte A, on voit que :

- en t_1 , Φ_1 précharge la capacité de structure si $E = 1$;
- en t_2 , Φ_2 la décharge si $E = 1$;
- en t_3 et t_4 , les informations sont disponibles.

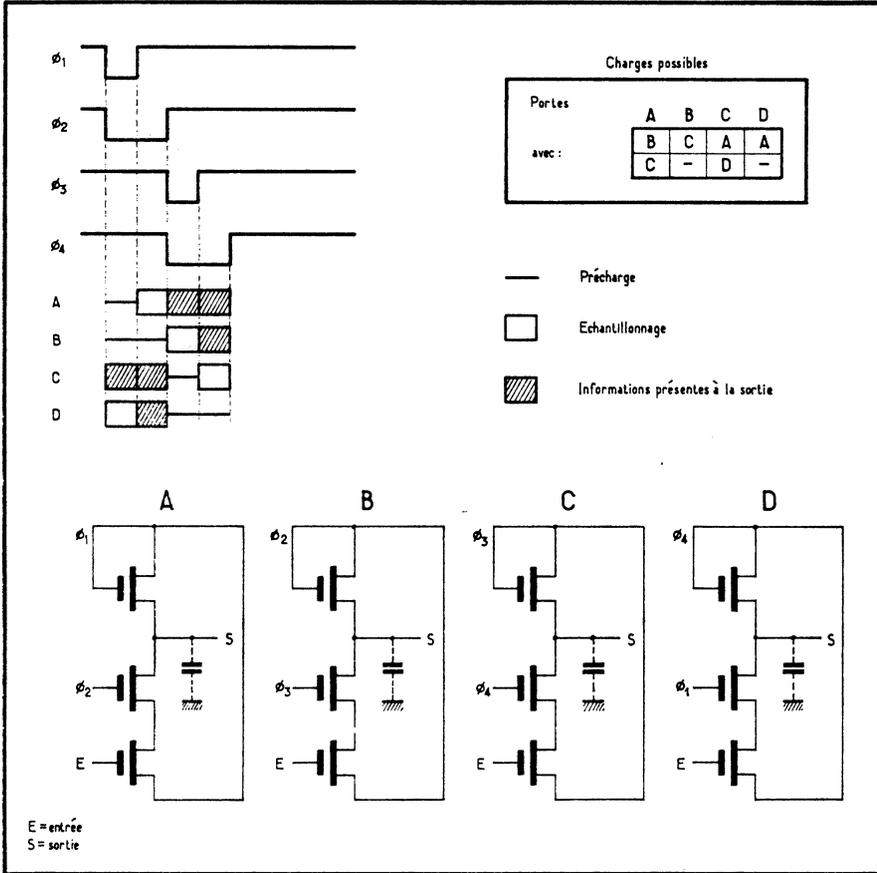


Fig. 20. — Distribution des quatre phases d'horloge lors de la constitution d'un registre dynamique de N bits, semblable à celui de la figure précédente. Les règles de charge, c'est-à-dire la succession possible des ordres, doivent être telles que le registre fonctionne correctement.

Les phases d'horloges sont à nouveau décalées, sans chevauchement, mais leur durée est ramenée à la normale et non doublée. Il est évident que l'information d'entrée utile ne doit être présente qu'en t_3 (pour A), faute de quoi le traitement n'aurait guère de sens.

7. REGISTRES DYNAMIQUES ALIMENTÉS PAR LES HORLOGES

L'étape ultérieure consiste à recourir à des cellules dynamiques non proportionnelles, mais alimentées par les horloges, selon le schéma de principe de la figure 19 :

- en t_1 , $\Phi_1 = \Phi_2 = 1$, par conséquent C_1 se charge au niveau 1, quel que soit le niveau de l'entrée, via Q_1 ;

- en t_2 , $\Phi_1 = 0$ alors que $\Phi_2 = 1$. Si l'entrée est à 1, la capacité C_1 se décharge via Q_2 et Q_3 . Si l'entrée est à 0, le transistor Q_3 est bloqué et le niveau de charge maintenu;
- en t_3 et t_4 , Φ_3 et Φ_4 agissent de même.

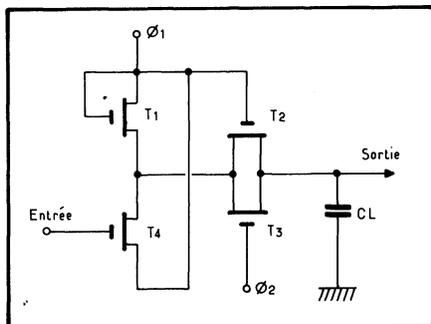


Fig. 21. — Étage non proportionnel alimenté par l'horloge.

Les règles de succession de telles portes sont indiquées figure 20. La figure 21 montre une autre disposition, à double MOS de transfert. Le nombre de structures possibles est, là aussi, important et reste généralement lié à la distribution des phases d'horloge (séparées, ou se chevauchant, de durées diverses...). Chaque système offre souvent des avantages qui lui sont propres.

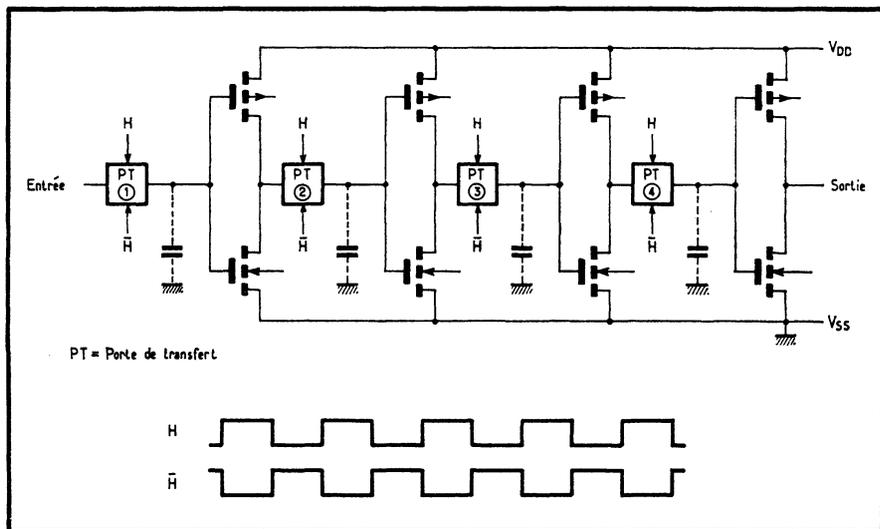


Fig. 22. — Registre dynamique à C/MOS. Les pavés notés PT sont des portes de transfert et reçoivent les signaux d'horloge en opposition.

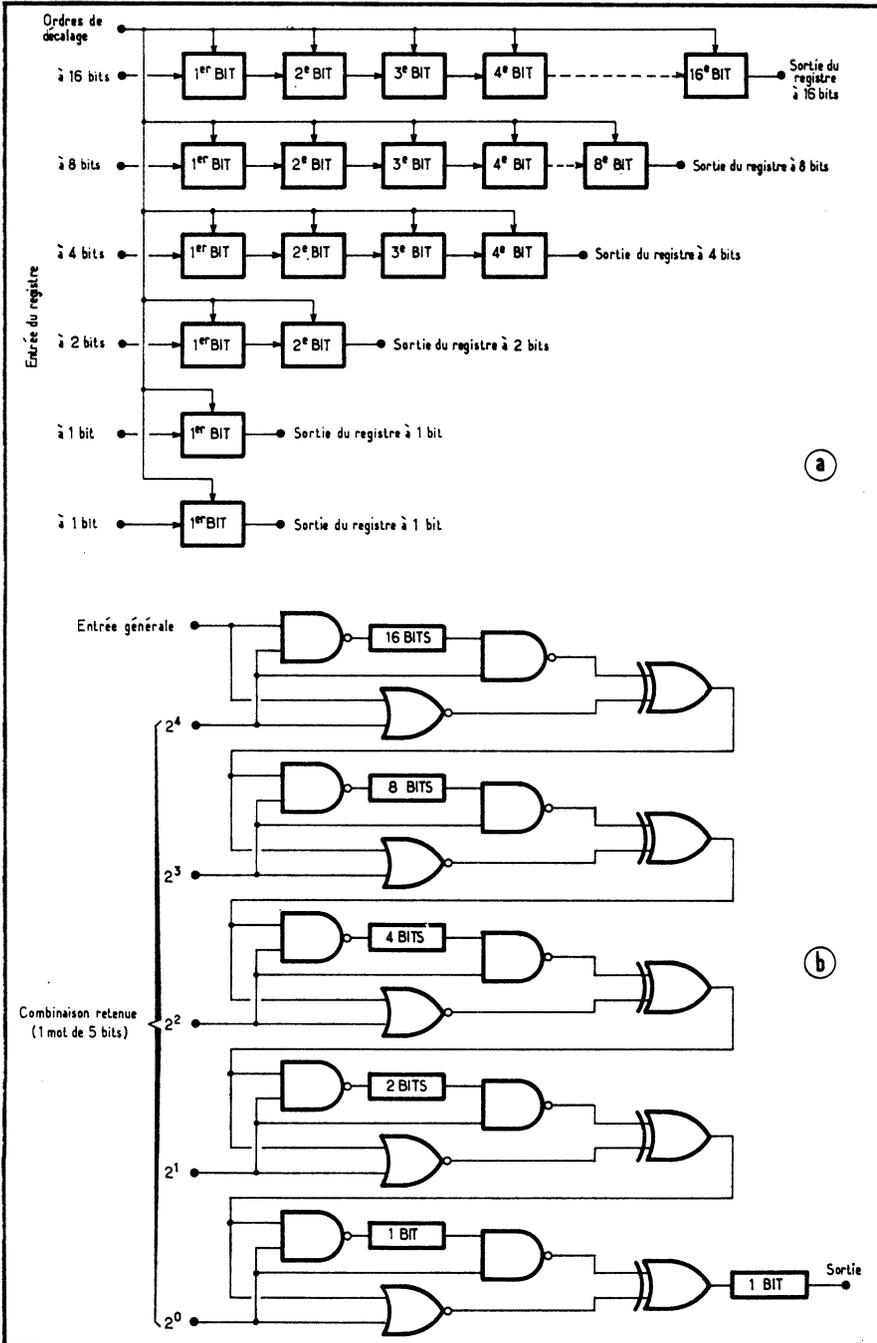


Fig. 23. — Le MEM 3032, dont le synoptique est donné en (a), est un registre à décalage programmable électriquement de 1 à 32 bits. Sa structure logique (b).

8. REGISTRES EN CMOS

Les registres en CMOS font appel aux mêmes principes, bascules ou inverseurs dynamiques. Dans ce dernier cas cependant, le schéma est très dépouillé puisque chaque inverseur n'est accompagné que par une porte de transfert (fig. 22).

9. REGISTRES A LONGUEUR AJUSTABLE

La longueur d'un registre, c'est-à-dire le nombre de points-mémoires en série qu'il comporte — ou encore, le nombre de bits qu'il peut emmagasiner — constitue l'une de ses caractéristiques fondamentales. Ainsi, on trouvera des registres à 8, 16, 32, 64 bits et jusqu'à plusieurs milliers de bits.

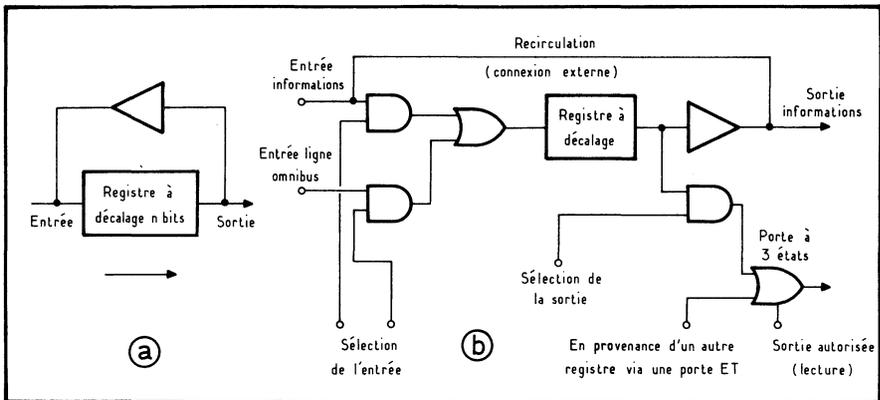


Fig. 24. — Principe du rebouclage d'un registre. Les informations circulent dans une boucle fermée.

Mais il est possible, dans un seul et unique CI, de loger plusieurs registres et même de les combiner entre eux, pour aboutir, dans le cas le plus complexe, à un registre de longueur ajustable électriquement.

Tel est le cas, par exemple, des FDN 126 et 136 de *RTC*, variables de 1 à 64 bits, ou du MEM 3032 de *General Instrument* que nous allons prendre pour exemple.

Ce CI comprend 6 registres, soit au total $16 + 8 + 4 + 2 + 1 + 1$ bit = 32 bits. Au lieu d'être indépendants, ils ont été reliés via des portes OU exclusif (fig. 23) qui reçoivent :

1. le signal de sortie du registre précédent;
2. un ordre, extrait d'un mot de 5 bits, qui sert à programmer les registres, via une porte OR.

En fonction de cette dernière information, le OU exclusif sélectionne et transmet, ou bien l'information d'entrée en court-circuitant le registre précédent, ou bien l'information à la sortie du registre précédent.

Le dernier registre de 1 bit restant toujours en service, on voit que la combinaison des autres permet toutes les valeurs de 1 à 32 bits.

10. LES REGISTRES REBOUCLÉS

Le registre à recirculation est un élément où la sortie est ramenée à l'entrée. Ainsi, une information d'entrée circule en permanence dans la boucle ainsi formée. Il se compose d'un registre à décalage proprement dit, et de portes logiques, comme le montre le synoptique de la figure 24.

CHAPITRE XIII

LES MOS COMPLÉMENTAIRES

L'étude des structures à MOS complémentaires a été abordée dans les chapitres précédents. En raison de l'importance qu'ont prise les familles à CMOS, il a cependant été jugé utile de leur consacrer un chapitre particulier. C'est donc une synthèse des CMOS qui est présentée ici, à partir des analyses faites antérieurement, mais développées et complétées sous un angle pratique.

1. UNE INTRODUCTION AUX CMOS

C'est RCA, qui, après avoir pendant de longues années étudié la logique à MOS complémentaires, a lancé la première série commerciale de CMOS sous le nom déposé de COS/MOS (*complementary symmetry/metal-oxyde-semiconductor*) : la série 4000. Celle-ci a fait école et a été suivie en tout premier lieu par Motorola avec les MCMOS (*Motorola CMOS*), puis par la plupart des fabricants.

National Semiconductor (NS) a réussi une percée vers les CMOS compatibles en brochage, fonctions et tensions avec la célèbre famille 54/74 TTL de *Texas*, tandis que diverses écoles technologiques tentaient d'améliorer les performances des CMOS. Parmi celles-ci apparaît en tout premier lieu la technologie des CMOS sur saphir, ou SOS (Silicon On Sapphire), ainsi que les CMOS à isolement diélectrique de *Harris Semiconductor*, sans omettre les divers procédés visant à la création des couches épaisses d'oxyde dans le substrat même, tel le LOCOS de *Philips* et *RTC*, ou l'*Isoplanar* de *Fairchild*, qui peuvent conférer aux CMOS des vitesses comparables à celles de la TTL.

En effet, le reproche fondamental adressé aux MOS est leur faible vitesse : 1 à 2 MHz pour les PMOS à l'origine, puis 5 à 10 MHz pour les NMOS qui suivirent. Or, avec des technologies améliorées, les CMOS prétendent aux vitesses de la TTL (10 ns de temps de propagation), et même de la TTL les plus rapide qui soit, la Schottky, avec un tpd de 3 ns.

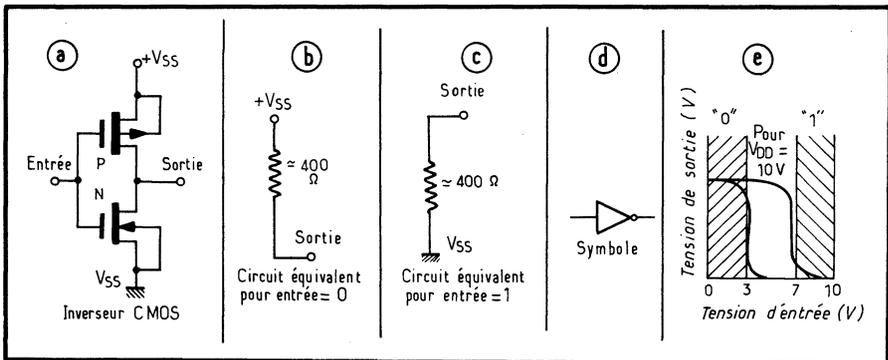


Fig. 1. — La porte CMOS.

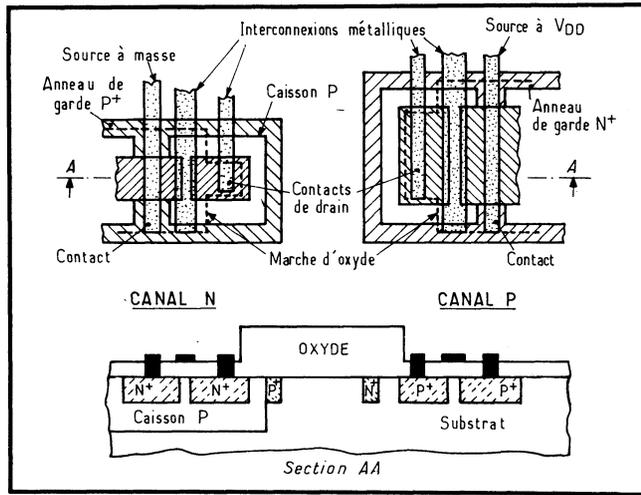
Par contre et sur ces familles, les CMOS offrent un avantage immense : une consommation qui, à fréquences basses ou relativement basses (moins de 3 à 10 MHz pour les CMOS classiques), peut être des centaines, des milliers ou même des millions de fois inférieure. Il en résulte bien sûr des possibilités d'applications nouvelles, avec tous les matériels portables ou autonomes, dont la montre électronique n'est qu'un exemple, mais encore des avantages immédiats sur des matériels classiques. Par exemple, tel terminal d'informatique exige un refroidissement énergique en fonctionnement car il dissipe trop de calories; s'il est réalisé en CMOS et non plus en TTL, on pourra éviter et éliminer le ventilateur et les circuits de ventilation. Ou encore, l'alimentation de tel appareil en logique TTL est volumineuse et chère; avec les CMOS, la puissance exigée pour la même fonction sera bien moindre et l'alimentation simplifiée d'autant; ce à quoi il faut ajouter que les découplages sont moins sévères avec les CMOS et que ceux-ci sont très peu sensibles aux variations de la tension d'alimentation.

Un second argument qu'on opposait aux CMOS, à l'origine, était leur prix. Or ce prix, au circuit, a connu une baisse constante, selon un processus caractéristique des circuits intégrés.

Enfin, on a beaucoup écrit sur la sensibilité des CMOS aux charges statiques, et leur propension à « claquer ». Or, là-aussi bien des progrès sont intervenus : désormais, toutes les entrées des CMOS sont protégées et peuvent supporter jusque 1 à 3 kV résultant de charges statiques. Malheureusement, certains environnements provoquent la naissance de tension atteignant 10 à 15 kV; en conséquence, et c'est là la rançon de leur grande impédance d'entrée, ils exigent un minimum de précautions lors de leur manipulation.

Ainsi, *les entrées ne devront jamais être laissées en l'air*; c'est la raison pour laquelle les fabricants livrent leurs CMOS montés sur des supports ou du plastique conducteurs. *On ne connectera rien à une broche de CMOS monté alors que l'alimentation est coupée* : par exemple, on ne branchera pas une pointe-test dont les capacités propres pourraient, en se déchargeant, endommager le circuit.

Fig. 2. — Plan et coupe d'un inverseur CMOS.



A plus forte raison, *les soudures s'effectueront avec un fer à souder relié à la masse*; l'opérateur qui manipule des CMOS devra, par surcroît de précaution, se relier lui-aussi à la masse, à partir d'un bracelet métallique (fil de câblage, par exemple), entourant son poignet. On recommande encore d'utiliser des *tapis de table en mousse conductrice*, et non isolante, si l'on devait y poser les CMOS, d'éviter les moquettes en nylon, etc.; toutes précautions qui, si elles n'étaient pas prises, n'entraîneraient pas fatalement des défaillances des circuits. Mais ne semblent-elles pas évidentes à la réflexion?

2. LA PORTE CMOS

La structure CMOS de base est l'inverseur. Puisqu'il a déjà été étudié dans le chapitre réservé aux inverseurs MOS de base, on va simplement ici en résumer les caractéristiques. Le schéma de l'inverseur est donné figure 1, avec ses montages

équivalents selon que l'entrée est attaquée par un 0 ou par un 1. Cependant, des anneaux de garde viennent, en réalité, entourer chaque MOS : anneaux N⁺ pour les PMOS, c'est-à-dire fortement dopés N, ou P⁺ pour les NMOS. Ils évitent toute fuite d'un élément à l'autre et servent aussi à amener le courant d'alimentation +V_{DD} via le substrat N, ou V_{SS} (la masse) via le substrat P (fig. 2).

Dans la série 4000 qui sert de référence, la tension d'alimentation peut prendre toute valeur V_{DD} entre 3 et 15 V, les éléments fonctionnant entre :

- -55 et +125 °C en boîtier céramique;
- -40 et +85 °C en boîtier plastique.

La dissipation est fonction de la fréquence F et donnée, si C est la capacité qui charge la sortie, par la formule simplifiée (fig. 3).

$$P \text{ dissipée} = C V_{DD}^2 F$$

Sous 5 V et hors des commutations (au repos), la consommation est réduite au seul courant de fuite et se mesure alors en microwatts ou en nanowatts.

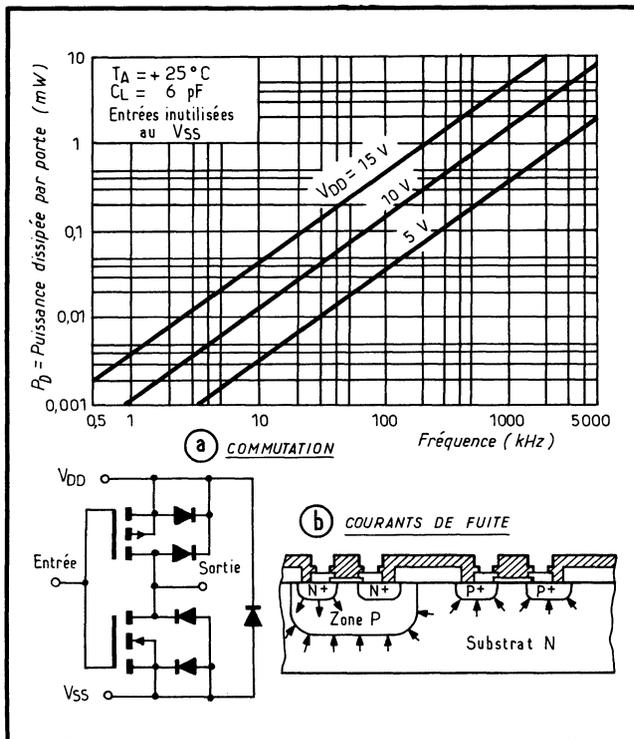


Fig. 3. — Consommation en fonction de la fréquence et courants de fuite des CMOS.

Les courants de commutation sont indiqués figure 4 en fonction de diverses valeurs de C et de temps de montée (t_r) et de chute (t_f). L'immunité au bruit, ainsi qu'on l'a définie précédemment, atteint 45 % de V_{DD}; c'est dire qu'elle est supérieure à celle de toutes les autres formes de logique, bipolaire incluses, à l'exception

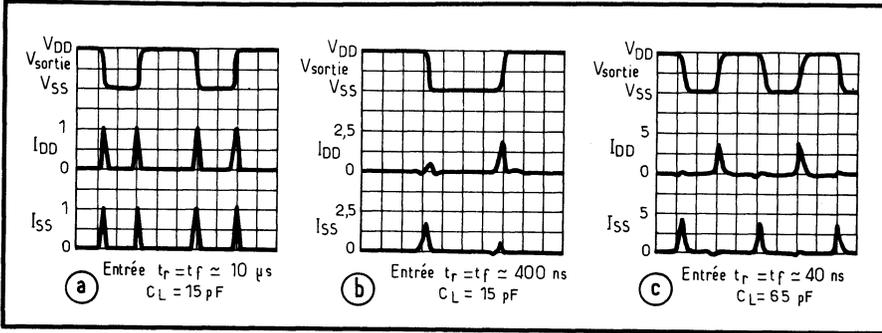


Fig. 4. — Les courants de commutation en fonction de diverses variables.

peut-être des familles spéciales à diodes Zener du type HNIL (*high noise immunity logic*).

Les entrées et sorties étant protégées par des diodes, on aboutit au schéma de la figure 5 pour l'inverseur fondamental. Il en résulte certaines obligations :

- Les signaux ne devront jamais être appliqués avant l'alimentation ou subsister après que l'alimentation soit coupée, pour éviter toute surcharge éventuelle de D_2 en service si l'alimentation est à la masse, ou toute action de redressement intempestif via D_2 si, l'alimentation étant supérieure, on applique un signal d'entrée alternatif.
- Les tensions d'alimentation ne devront pas être interverties, ce qui semble aller de soi! On devra toujours avoir, pour le moins :

$$V_{DD} - V_{SS} > -0,5 \text{ V}$$

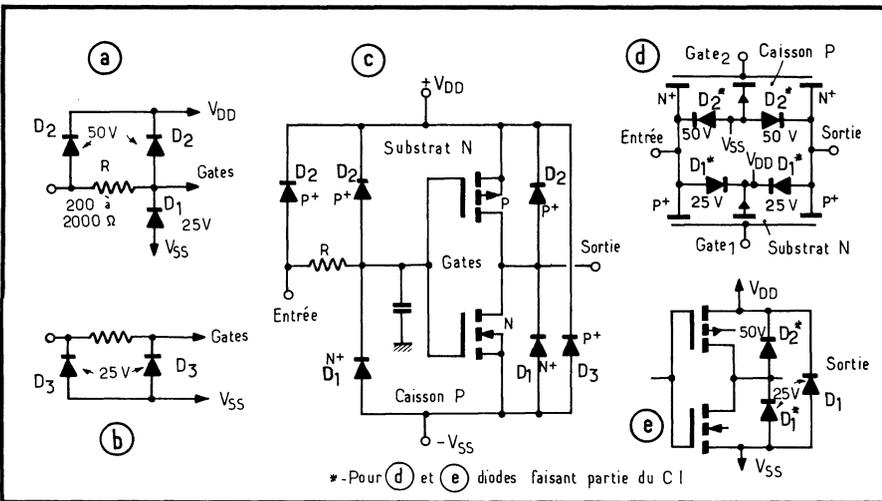


Fig. 5. — Les protections d'entrées-sorties.

- Les signaux de commande devront être compris entre V_{DD} et V_{SS} .
- Les CI ne seront pas insérés ou extraits de leur support alors que les tensions sont appliquées, pour exclure tout risque de transitoires.
- Toutes les entrées inutilisées seront reliées au V_{SS} ou au V_{DD} selon les besoins logiques du montage.
- Dans l'ordre des soudures, la connexion au V_{DD} précédera celle au V_{SS} .

Remarque à propos de la consommation

Les courbes classiques comparant la consommation des CMOS en fonction de la fréquence à celle d'autres familles, TTL par exemple (fig. 6), montrent effectivement l'économie que procurent les CMOS aux faibles fréquences. Mais l'avantage semble se perdre dès qu'on monte un tant soit peu en fréquence, et entre 2 et 10 MHz, les courbes de consommation se recoupent généralement. Il semblerait alors qu'il n'y ait plus intérêt à recourir aux MOS complémentaires. Or, une telle conclusion se révélera souvent erronée car :

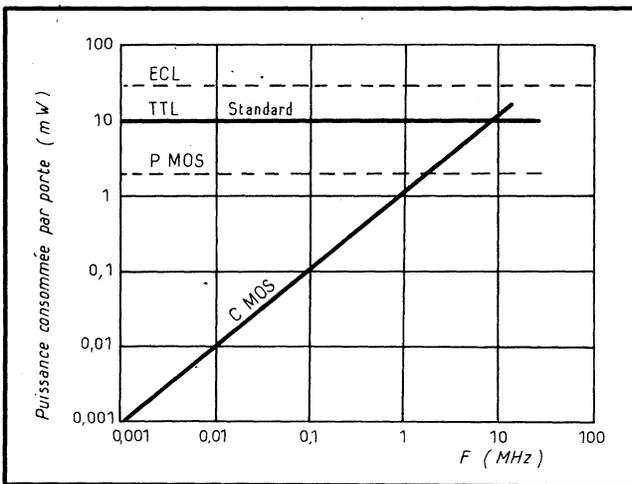


Fig. 6.— Courbes habituelles montrant l'accroissement de la consommation des CMOS en fonction de la fréquence. Elles sont discutées dans le texte.

1. La charge capacitive d'une porte est, le plus souvent, constituée par l'entrée de la porte suivante intégrée sur la même puce, donc très faible (moins de 5 pF). Ce n'est qu'en fin de chaîne et à la sortie du CI qu'on trouvera, par exemple, une charge supérieure.

2. La porte qui ne commute pas ne consomme pas... contrairement à ce qui se passe avec les autres logiques classiques.

3. Dans une chaîne de division, par exemple, seul l'étage d'entrée travaille à la plus haute fréquence, mais pas les autres circuits (cas de la montre à quartz!). Or, la consommation est liée à la fréquence.

4. Dans un décodeur, à 1 parmi 16, par exemple, seuls deux portes sur 16 commutent pour une impulsion d'entrée : celle qui s'ouvre, et celle qui se ferme;

les autres restent au repos et ne consomment donc pas; il en va de même pour bien d'autres circuits.

Ces remarques montrent que le calcul de la consommation ne peut se faire linéairement, en extrapolant à partir de ces courbes. Une excellente preuve de ce fait était fournie en 1974 par la société *CGCT* qui réalisait un central téléphonique expérimental, de 16 000 voies et 5 000 erlangs, capable de desservir quelque 80 000 abonnés; bien que fonctionnant à 2 MHz à impulsions codées (MIC), avec des impulsions espacées de 488 ns, ce central qui a été réalisé en une seule baie ne consomme que 800 W, grâce aux CMOS; il regroupe quelque 5 000 CI dont 4 500 standards. Hors, un calcul préalable et sans nuances de la consommation aurait mené à des valeurs telles que les CMOS auraient probablement été abandonnés. Grâce à la faible consommation réelle et pratique, une seule armoire a ainsi suffi là où, en raison de la puissance dissipée, il en aurait fallu plusieurs interconnectées dans des conditions difficiles (les horloges auraient dû être synchronisées à mieux que 50 ns; or, en 50 ns, un signal parcourt 8 m : les câbles doivent donc être beaucoup plus courts...). Ajoutons que ce central sera — peut-être — opérationnel vers les années 80.

La mise en parallèle des entrées

Si l'on relie en parallèle plusieurs entrées d'un NOR ou d'un NAND multiple (fig. 7 a), on accroît sensiblement la vitesse de travail car la résistance équivalente des canaux conducteurs, en parallèle, est réduite. Ce gain est cependant tempéré par l'augmentation de la capacité d'entrée.

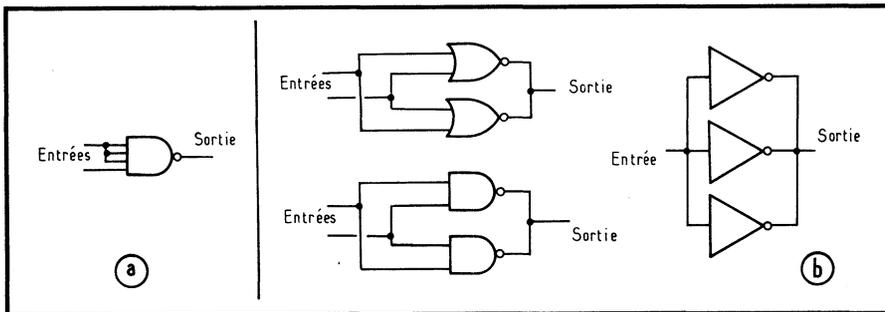


Fig. 7. — Mise en parallèle des entrées ou des portes CMOS.

La mise en parallèle de portes, elle, sert à accroître le courant de sortie, absorbé ou émis selon que c'est le PMOS ou le NMOS qui est conducteur (fig. 7 b). C'est là un moyen d'accroître la sortance.

L'immunité au bruit

L'immunité au bruit *typique* est de 45 % de la tension d'alimentation V_{DD} ; l'immunité *garantie*, lorsque les conditions les plus défavorables sont réunies, est

Tableau A. — Immunités au bruit comparées

FAMILLE (porte type)	ALIMENTATION (V)	CONSUMMATION AU REPOS (mW)	TEMPS DE PROPAGATION (ns)		IMMUNITÉ EN TENSION CONTINUE - SIGNAL (V)				IMMUNITÉ EN ALTER - NATIF		IMPÉDANCE DES LIGNES - SIGNAL (Ω)		IMMUNITÉ TYPIQUE EN ÉNERGIE (nJ) POUR UNE DURÉE D'IMPULSION Dimp (ns)			
					V _{NL}		V _{NH}		ALIMENT. (V)				ÉTATS:			
					Min	Typ	Min	Typ	+	Mosse					L	H
					t _{PHL}	t _{PLH}	Min	Typ	Min	Typ			E _{NL} (nJ)	Dimp (ns)	E _{NL} (nJ)	Dimp (ns)
DTL (MC 849)	5	8	20	50	0,7	1,2	0,7	3,8	3	1	49	1,8k	1,4	45	0,4	40
TTL (MC 7400)	5	10	8	12	0,4	1,2	0,4	2,2	3	1	30	140	1,7	20	1	25
HTL (MC 672)	15	25	85	130	5	7,5	4	7	6	4,5	140	1,6k	60	125	5	145
CMOS (MC 14011)	5	25.10 ⁻⁶	35	100	1,5	2,2	1,5	3,4	2,8	1	1,7k	4,8k	1	155	0,9	280
	10	50.10 ⁻⁶	20	35	3	4,2	3	6	5,7	4,3	670	1,5k	3,7	70	3,1	90
	15	150.10 ⁻⁶	8	15	4,5	6,3	4,5	9	8,5	6,4	460	1k	7,2	50	8,5	75

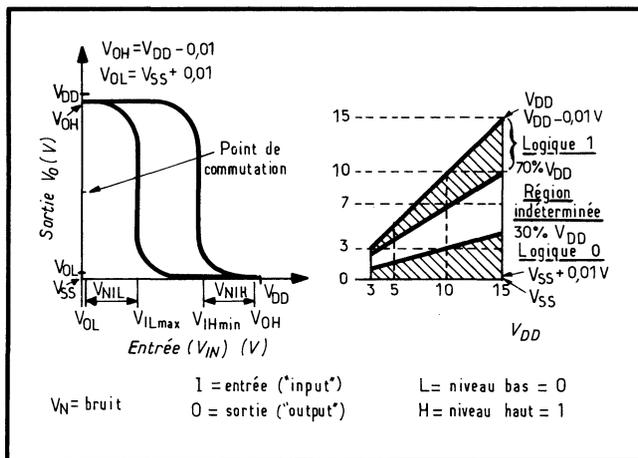


Fig. 8. — Fonction de transfert de l'inverseur et valeurs garanties de l'immunité au bruit.

de 30 % de V_{DD} (fig. 8). Si $V_{DD} = 5 V$, on peut définir les conditions logiques pour deux portes, reliées sortie à entrée (fig. 9).

Le tableau A compare l'immunité du bruit de la série *Motorola* 14000 à celle des familles bipolaires de la même firme, en tension (signal et alimentation) et en énergie. On voit que les CMOS sont ici supérieurs à la DTL et TTL 54/74, mais moins efficaces que la HNIL (logique bipolaire à haute immunité, à Zener), laquelle est d'ailleurs fort lente.

Fig. 9. — Caractéristiques d'entrée-sortie.

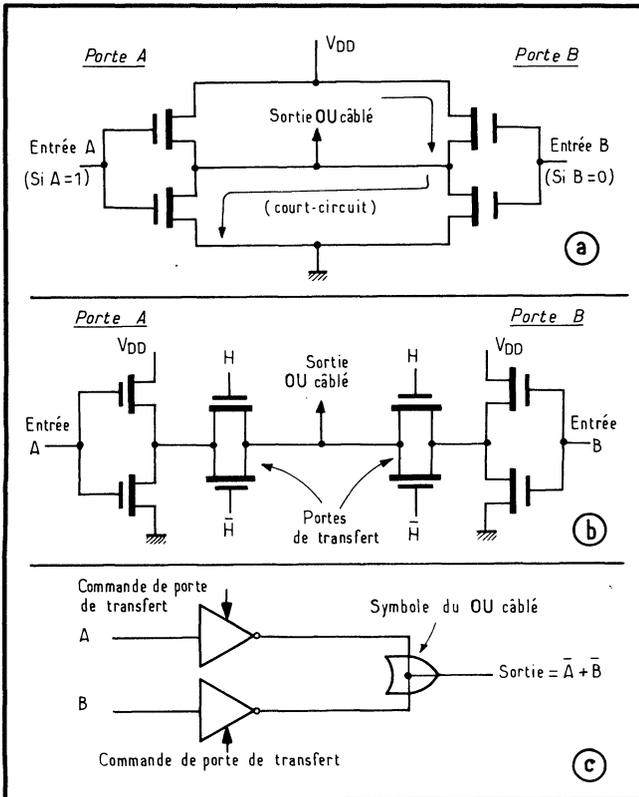
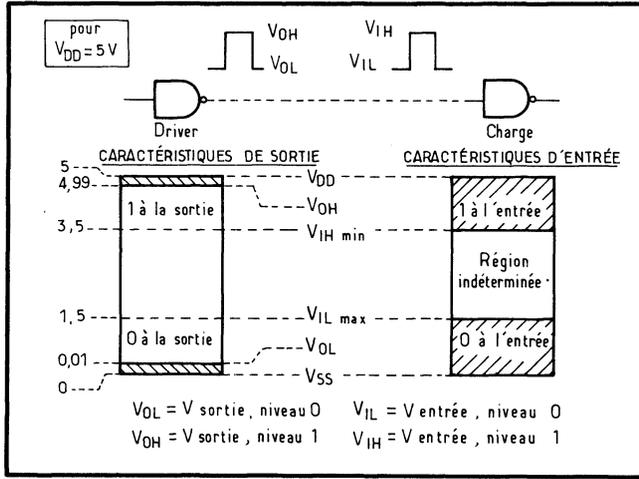


Fig. 10. — Impossible directement, le OU câblé (a) est permis grâce aux portes de transfert (b). Symbole (c).

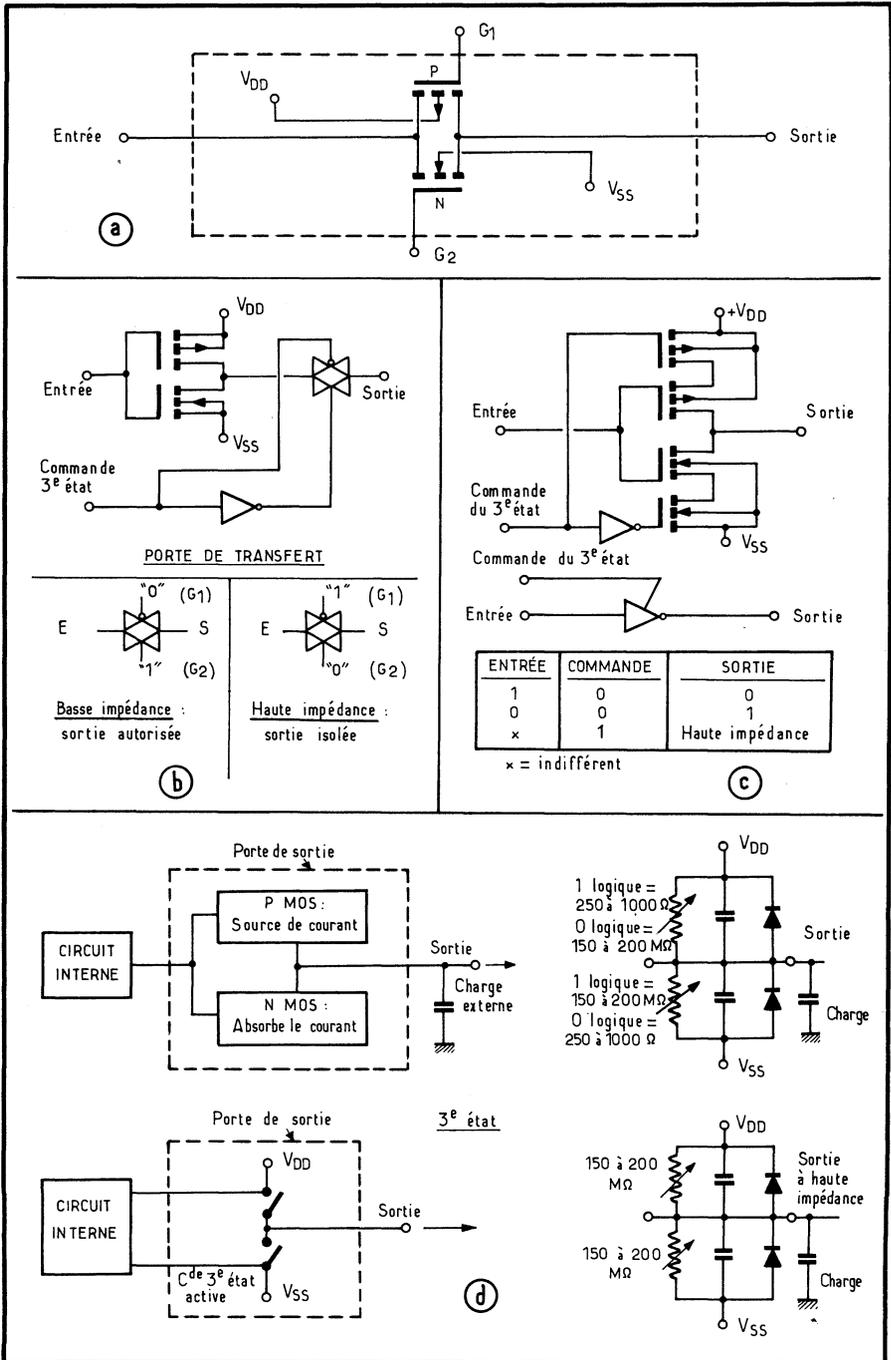


Fig. 11. — La porte de transfert (a) et l'un de ses symboles (b), selon Motorola. Autre version de commande 3^e état (c). Représentation équivalente des sorties (d).

Le OU câblé et la logique 3 états

Comme pour la TTL, on ne peut relier entre elles les sorties d'inverseurs CMOS pour réaliser le OU qu'on dit alors « câblé », ou encore « fantôme ». En effet, on risque un court-circuit pur et simple entre les tensions d'alimentation lors que les deux portes ne sont pas dans le même état (fig. 10 a).

Par contre, il est très facile de transformer en « 3 états » les circuits CMOS puisqu'il suffit alors d'introduire une porte de transfert en série sur la sortie. Lorsqu'elle est conductrice, la sortie travaille normalement, c'est-à-dire qu'elle passe aux niveaux 1 ou 0; mais lorsqu'elle est bloquée, la sortie apparaît « en l'air », niveau indéfini à haute impédance qu'on appelle *troisième état* (fig. 10 b). Le schéma de la porte de transfert et l'un de ses multiples symboles (Motorola : deux portes tête-bêche imbriquées) sont donnés figure 11 a. Une autre méthode aboutissant à créer ce même 3^e état est indiquée en c : elle consiste à introduite deux MOS en série avec un inverseur; l'action est identique. Enfin, et selon que la sortie est au niveau 0, au niveau 1 ou isolée, on obtient les schémas équivalents donnés en d. Grâce à ce 3^e état et en raison des très faibles courants de fuite, plus de 50 portes CMOS peuvent être réunies en OU câblé.

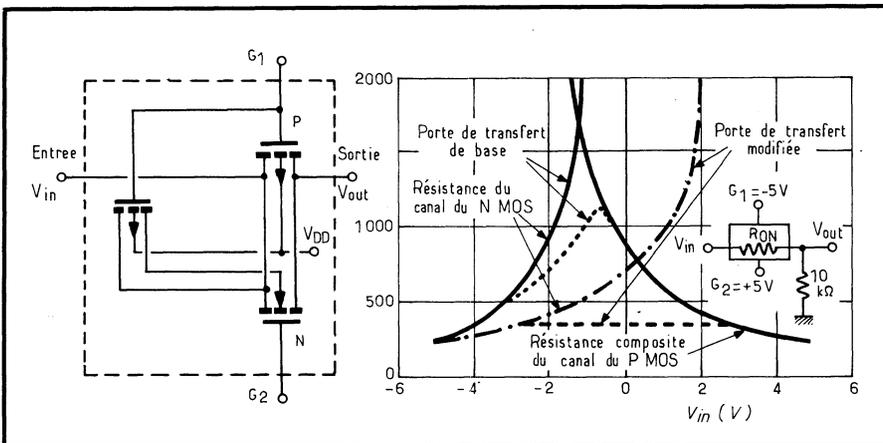


Fig. 12. — Porte de transfert modifiée et courbes résultantes.

On notera encore que la porte de transmission peut revêtir d'autres aspects, comme le montre la figure 12 : un transistor supplémentaire commande la polarisation du substrat du NMOS. De ce fait, le NMOS demande un temps plus long pour se bloquer, mais il en résulte une variation bien plus plate de R_{on} en fonction de la tension d'entrée. Un tel montage a été adopté dans le quadruple commutateur analogique MC 14016, par exemple; la résistance R_{on} varie alors de +15 % seulement autour de sa valeur nominale pour toute la gamme des tensions d'entrée et sous 10 V d'alimentation.

On verra, ultérieurement, que des structures encore plus complexes peuvent être sélectionnées.

Logiques positive et négative

Une certaine confusion peut résulter de l'usage de la logique positive avec les CMOS, alors que les familles MOS monocanales P s'inspirent de la logique négative. Avec un peu de réflexion, toutefois, on va voir que la conversion est excessivement simple.

Rappelons que :

- *en logique positive*, la tension la plus élevée est le 1 (par exemple, le +5 V est le 1 et le 0 V le zéro);

- *en logique négative*, la tension la plus basse négativement est le 1 (par exemple : -5 V est le 1 et la masse, ou 0 V, le zéro).

Ainsi, en règle générale, le niveau zéro de référence reste la masse. Si l'on appelle L (du *low* américain) le niveau *bas*, c'est-à-dire le niveau de tension le plus

Tableau B. — Logiques positive et négative comparées

TABLE COMMUNE			LOGIQUE POSITIVE L=0 et H=1 Fonction NOR			LOGIQUE NÉGATIVE L=1 et H=0 Fonction NAND		
X	Y	Sortie	X	Y	Sortie	X	Y	Sortie
L	L	H	0	0	1	1	1	0
L	H	L	0	1	0	1	0	1
H	L	L	1	0	0	0	1	1
H	H	L	1	1	0	0	0	1
L = niveau bas H = niveau haut			Sortie = $\bar{X} \cdot \bar{Y}$ = $\bar{X + Y}$			Sortie = $\bar{X} + \bar{Y}$ = $\bar{X \cdot Y}$		

bas (0 en logique positive, et 1 en négative), et H (de *high*) le niveau le plus *haut* (1 en logique positive et 0 en logique négative), on peut dresser les tables d'états B où l'on trouve les états du NOR en logique positive, équivalent au NAND en logique négative, et leur table commune où les niveaux sont notés L et H. Ainsi :

- un NOR en CMOS, en logique positive, devient un NAND en logique négative (ce qu'enseigne l'algèbre de Boole);
- le NAND en CMOS en logique positive devient un NOR en logique négative, pour les mêmes raisons;
- un compteur *comptant* en logique positive devient *décomptant* en logique négative; les entrées R à Z et R à 1 s'inversent elles aussi (*SET* devient *RESET*, et vice versa). Il est facile de retrouver un montage *comptant* en inversant les niveaux d'entrée et de sortie tout simplement;
- on pourrait continuer dans le même ordre d'idées à examiner ce que deviennent les autres types de circuits.

3. L'INTERFAÇAGE DES CMOS

Pour associer les CMOS à d'autres types de logique (DTL, TTL, ECL, MOS monocanaux...), il convient de respecter les caractéristiques de niveaux logiques et de tensions d'alimentation pour garantir, à chaque fois, les niveaux de dissipation admis et l'immunité au bruit.

Le courant d'entrée des CMOS, qui n'est qu'un courant de charge de quelque 5 pF, est négligeable. Le courant de sortie, lui, reste faible si on le compare à celui des logiques bipolaires : de l'ordre de 0,5 mA (0,3 à 1 mA généralement). Le MOS à canal N *absorbe* un courant tandis que le MOS à canal P *émet* un courant.

TTL à CMOS

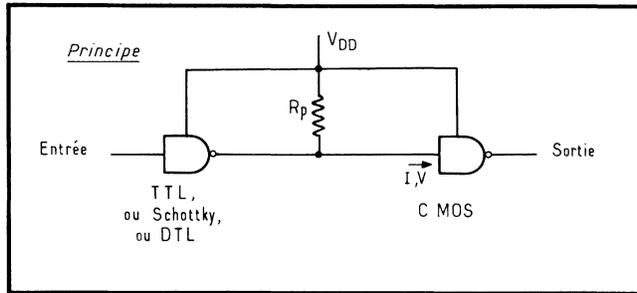
Pour une tension d'alimentation commune de +5 V, la TTL fournit à la sortie :

- 0,4 V max. (pour 10 pA) au 0;
- 3,6 V min. au 1.

Si l'on se réfère à la COS/MOS 4000, celle-ci accepte à l'entrée :

- 1,5 V max. et 10 pA pour une commutation garantie au niveau 0;
- 3,5 V min. et 10 pA au 1.

Fig. 13. — L'interfaçage TTL à CMOS.



De ce fait :

- au 0, la TTL pilote directement la CMOS avec une immunité de $1,5 - 0,4 = 1,1$ V. Celle-ci est donc sensiblement inférieure à l'immunité d'une porte CMOS en commandant une autre, la marge d'immunité passant de 30 % à 22 %;
- au niveau 1, par contre, l'immunité est par trop réduite : $3,6 - 3,5 = 0,1$ V. C'est pourquoi l'on utilise une résistance de rappel notée R_p (fig. 13), qui porte à V_{DD} le niveau 1 de sortie TTL. Sa valeur optimale dépend du montage considéré.

Les meilleures valeurs seront calculées en utilisant les circuits TTL à sortie « à collecteur ouvert » et non plus en totem-pôle.

CMOS à TTL

Dans ce cas, inverse du précédent, on trouve à la sortie de la porte CMOS (série 4000) :

- 4,4 V min. pour 4,5 V d'alimentation au 1 logique; or, l'alimentation est normalement de 5 V. L'entrée TTL, elle, se contente de 2,9 V min. pour un courant de 40 μ A max. Pratiquement, l'immunité au bruit est ici de l'ordre de 50 %, donc excellente;
- au niveau 0, la sortie CMOS fournit 0,8 V max. pour un courant absorbé (car c'est le MOS relié à la masse qui conduit, donc il absorbe un courant positif, venant de l'extérieur) de 1,8 mA : tel est le cas, par exemple, de la porte type CD 4001. Pour les buffers CD 4009 et 4010, le courant absorbé serait de 7 mA sous 0,8 V. Or, l'entrée TTL exige 0,8 V max. et un courant de 1,6 mA pour la TTL standard, c'est-à-dire 54/74. Malheureusement, l'immunité en bruit à 0,8 V est quasiment inexistante.

Ici, il faut donc réduire la tension de sortie de 0,4 V pour rétablir cette immunité; à ce moment, par contre, le courant absorbé est réduit et seules, les portes de puissance type CD 4009 et 4010 sont capables d'absorber à ce niveau plus de 1,6 mA (exactement 3 mA). A ce compte, la liaison CMOS à TTL est directe (fig. 14 a et b).

Une autre méthode consiste à connecter en parallèle des portes multiples, comme on l'a évoqué antérieurement, pour accroître le courant absorbé en sortie au niveau *bas* (fig. 14 c et d). Il est aussi facile de relier directement la CMOS à la TTL, en particulier la Schottky aujourd'hui, mais il sera bon de se référer à chaque fois aux valeurs de courant annoncées par les fabricants. Bien entendu, l'usage de la TTL faible puissance, ou plus exactement de la Schottky faible puissance, se révèle ici des plus intéressants.

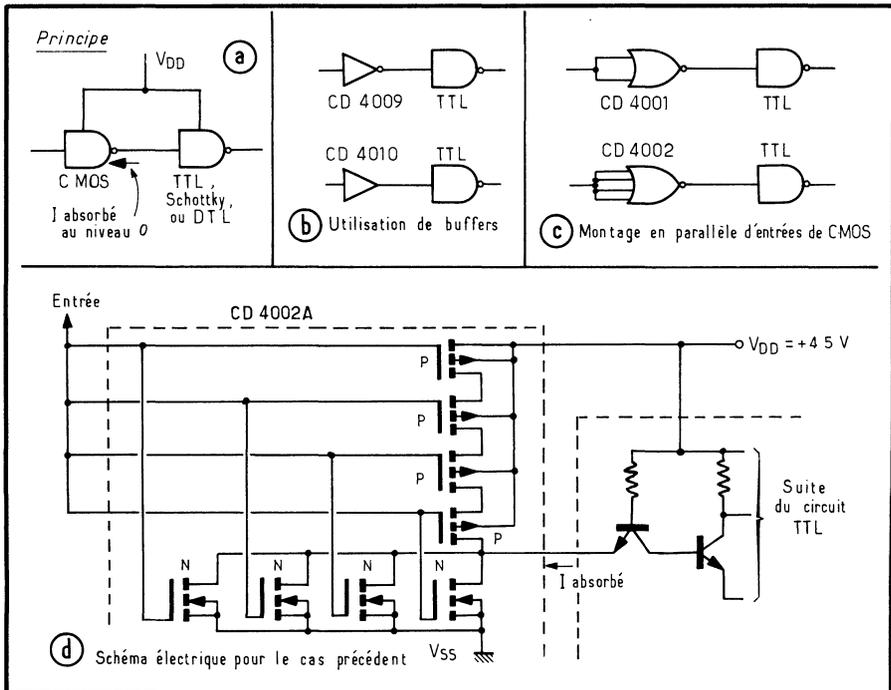


Fig. 14. — Interfaçage CMOS à TTL.

Le décalage des niveaux

Si les circuits CMOS travaillent sous +15 V ou toute autre tension supérieure à 5 V, il faut assurer un décalage des niveaux. On l'obtient très simplement à l'aide de ces mêmes buffers 4009 et 410, dotés par construction de deux lignes distinctes d'alimentation (fig. 15).

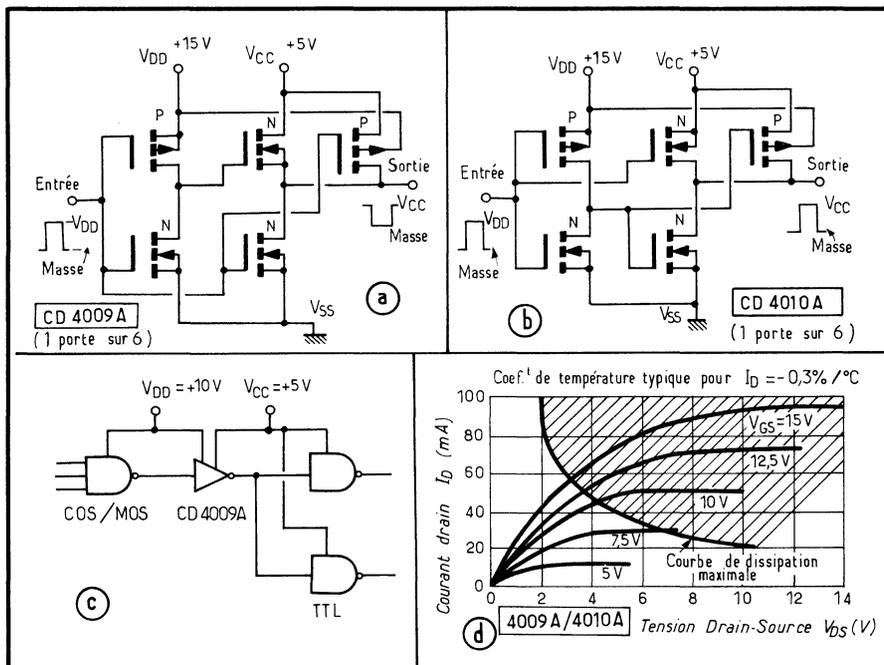


Fig. 15. — Schémas électriques des buffers 4009 (a) et 410 (b). Utilisation en conversion de niveau (c). Dissipation maximale (d) de ces mêmes buffers.

La tension V_{DD} sera ici le +15 V, par exemple, et le V_{CC} le +5 V alimentant également la TTL. L'avantage revient, pour les CMOS en amont, à pouvoir fonctionner sous une tension plus élevée, donc à vitesse supérieure.

CMOS et HTL

La HTL, ou les familles du même type à nom dérivé, sont des logiques bipolaires DTL ou TTL dotées de diodes Zener qui en accroissent les niveaux et, partant, l'immunité au bruit, ce qui est l'objectif recherché. En général, ces logiques fonctionnent sous +15 V avec une immunité de 45 %.

Tout dépend ici de la famille réelle considérée et de la structure de la porte. En général, les liaisons seront directes et, dans le sens CMOS à HTL, on retrouvera le problème du courant absorbé par la CMOS au niveau bas (fig. 16); pour la meilleure immunité, on utilisera à nouveau les buffers.

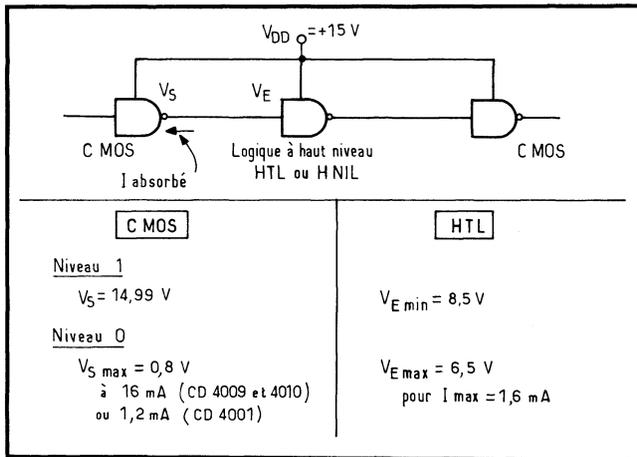


Fig. 16. — Interfaçage CMOS-HTL ou HNIL-CMOS.

Tableau C. — Interfaçage des CMOS

INTERFACE	IMMUNITÉ AU BRUIT (V)		NIVEAUX LOGIQUES (V)		SORTANCE MAX.	NOTES
	1	0	1	0		
	CMOS à CMOS	1,5	1,5	3,5		
CMOS à HTL (1)	3	3	7	3	> 50	Sous 10 V
	4,5	4,5	10,5	4,5	> 50	Sous 15 V
HTL à CMOS	5	5	8,5	6,5	1	
CMOS à L/TTL	3	3	10,5	4,5	> 50	Charge active HTL Charge 2 kΩ à 5 kΩ HTL
	4	4				
L/TTL à CMOS	2,5	0,4	2	0,7	1	Pour accroître la sortance, connecter des portes CMOS en parallèle.
CMOS à TTL	1,1	1,2	3,5	1,5	—	
TTL à CMOS	2,5	0,4	2	0,8	2	Avec buffers CMOS.
CMOS à PMOS	1,1	1,1	3,5	1,5	—	
	3	4	-3	-9	> 50	PMOS haut seuil.
PMOS à CMOS	2,5	6	2,5	1	> 50	PMOS faible seuil.
	3,9	3,9	3,5	1,5	> 50	PMOS haut seuil.
CMOS à ECL	3	3	2	-2	> 50	PMOS faible seuil.
	0,225	4,325	-1,05	-1,425	2	$V_{SS} = -5,2 \text{ V}$ et $V_{DD} = 0$
ECL à CMOS	0,66	1,56	-1,56	-3,64	> 50	$V_{SS} = -5,2 \text{ V}$ et $V_{DD} = 0$

(1). Bipolaire à haute immunité à Zener.

(d'après MOTOROLA)

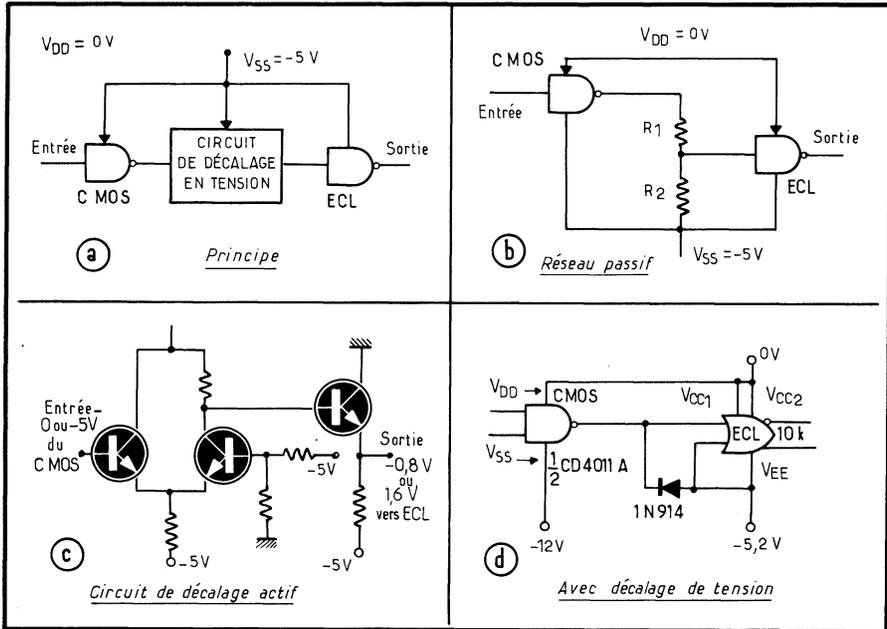
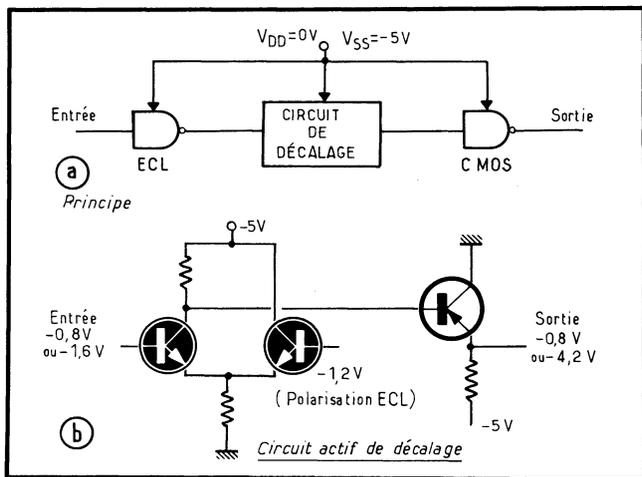


Fig. 17. — Interfaçage CMOS à ECL.

CMOS à ECL

La logique la plus rapide qui soit est, en 1978, l'ECL, logique bipolaire non saturée à couplage par les émetteurs : elle dépasse le gigahertz. Par contre, l'excur-

Fig. 18. — Interfaçage ECL à CMOS.



sion logique en tension est faible; ainsi, les niveaux seront de $-0,8\text{ V}$ max. pour le 0 et $-1,6\text{ V}$ min. pour le 1, l'alimentation étant assurée sous -5 V (ou $-5,2\text{ V}$).

Tout d'abord, il faudra alimenter le CMOS sous -5 V et faire alors $V_{DD} = 0\text{ V}$ et $V_{SS} = -5\text{ V}$ (V_{SS} était antérieurement la masse). Puis, le CMOS fournissant les niveaux logiques 0 et -5 V , on réalisera des circuits de décalage (fig. 17), soit :

- *Passif*, à diviseur résistif; ce système est très sensible à la température et dérive aisément; d'autre part, sa constante de temps en fait un circuit lent;
- *Actif*, à transistors bipolaires, qui exclut les deux défauts précédents.

Il est possible de relier directement un CMOS à un circuit ECL du type 10 K, par exemple, en introduisant une diodé limitant l'excursion négative (fig. 17 d).

ECL à CMOS

C'est là le problème inverse mais si le niveau 1 de $-0,8\text{ V}$, à la sortie de l'ECL,

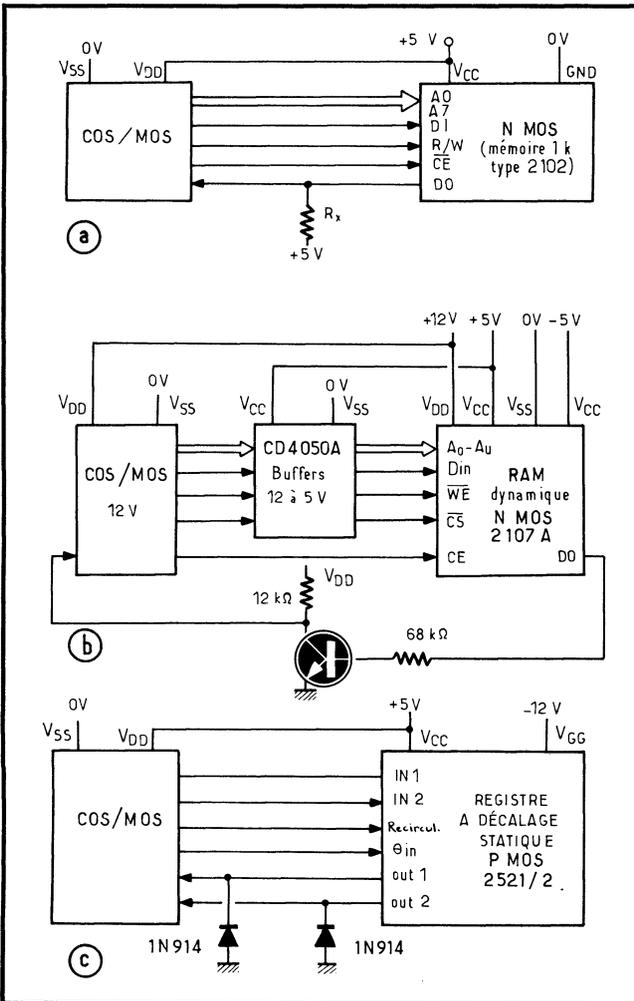


Fig. 19. — Interfaçage CMOS à NMOS, appliqué à une mémoire 1 K type 2102 (a) et à une RAM dynamique 2107 avec buffers (b). Interfaçage CMOS à PMOS appliqué à un registre statique (c).

convient à l'entrée du CMOS, il n'en va pas de même du niveau 0 de $-1,6\text{ V}$ (sortie ECL) qu'il faut porter à $-4,2\text{ V}$ pour le CMOS. L'interface est ici à transistors bipolaires (fig. 18).

CMOS et MOS monocanaux

Les tensions d'alimentation et les niveaux logiques des CMOS et des MOS monocanaux sont compatibles; de ce fait, les liaisons peuvent être directes. Cependant, on distinguera deux cas :

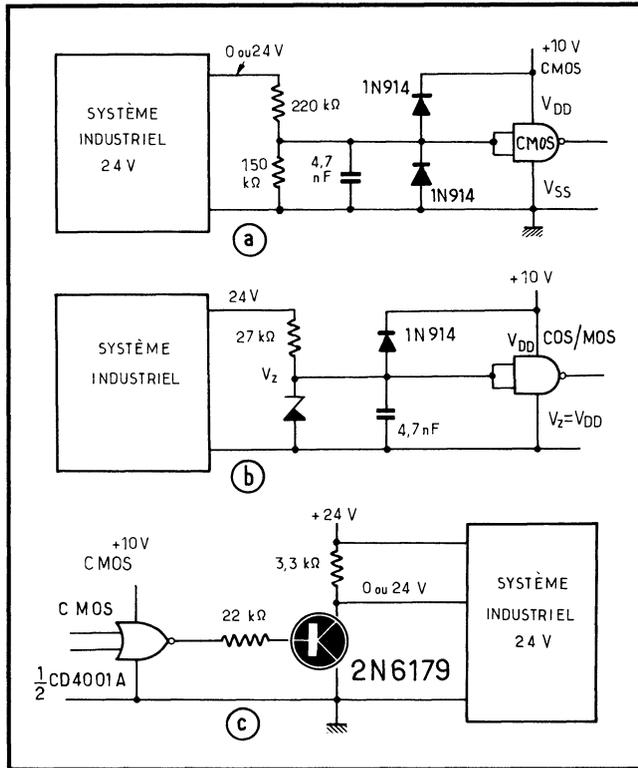


Fig. 20. — Circuits CMOS commandés par une logique industrielle (a) de 24 V, et variante à Zener (b), CMOS commandant un système industriel 24 V (c).

1. MOS monocanaux canal N : la tension d'alimentation est positive et l'on utilise généralement la logique positive. Il n'y a donc aucun problème.
2. MOS monocanaux canal P : ici, l'alimentation est généralement négative pour une logique négative. Il convient donc de transposer les logiques, comme on l'a montré précédemment.

Le tableau C résume quelques aspects des problèmes d'interfaçage, ici, et la figure 19 donne quelques structures possibles pour divers cas.

CMOS à systèmes industriels et à d'autres composants

Les montages de la figure 20 montrent comment le CMOS est attaqué, ou attaque un système industriel fonctionnant sous 24 V. En *a*, on transpose le 24 V en niveau 10 V; en *b*, une Zener remplace la diode normale; dans les deux cas, la présence du condensateur garantit une excellente immunité au bruit. En *c*, un transistor actif sert d'interface.

Figure 21, le CMOS attaque un relais (*a*) ou un solénoïde actionnant un marteau, et ce via un transistor (*b*). Il commanderait directement un thyristor

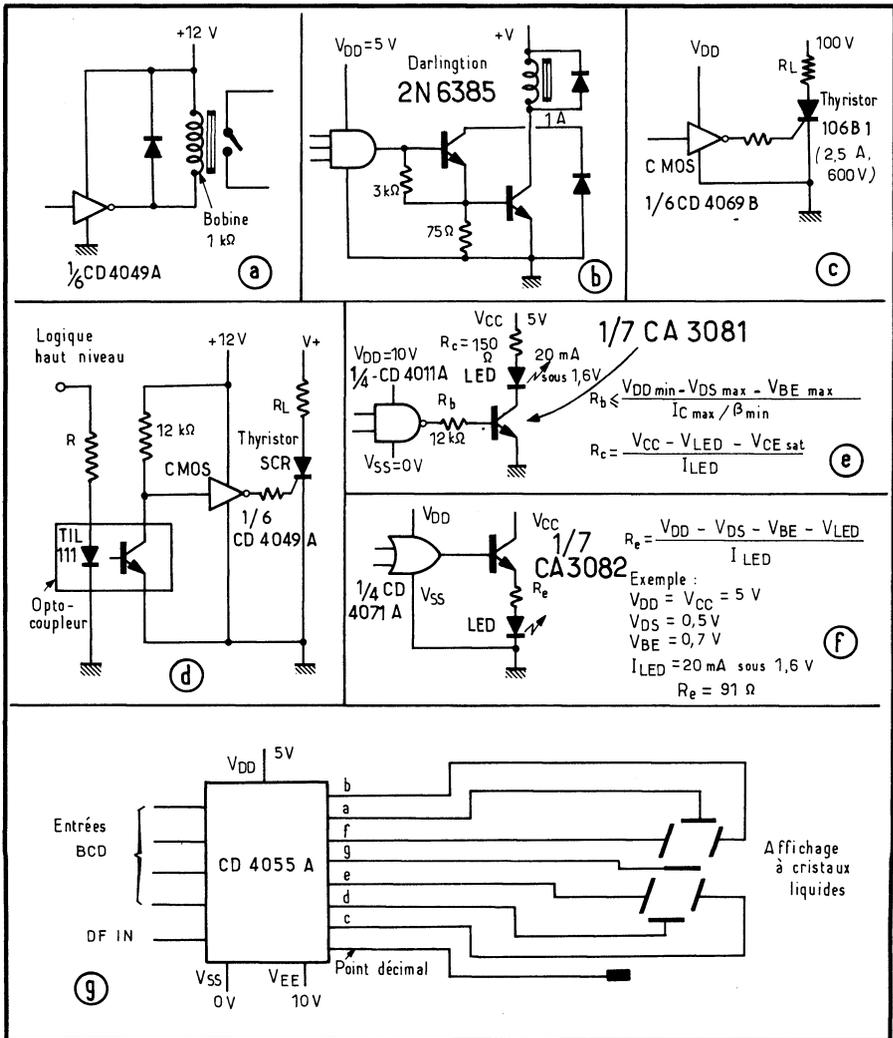


Fig. 21. — CMOS commandant un relais (a), un marteau d'imprimante (b), un thyristor (c et d), une LED (e et f) ou un affichage à LCD (g).

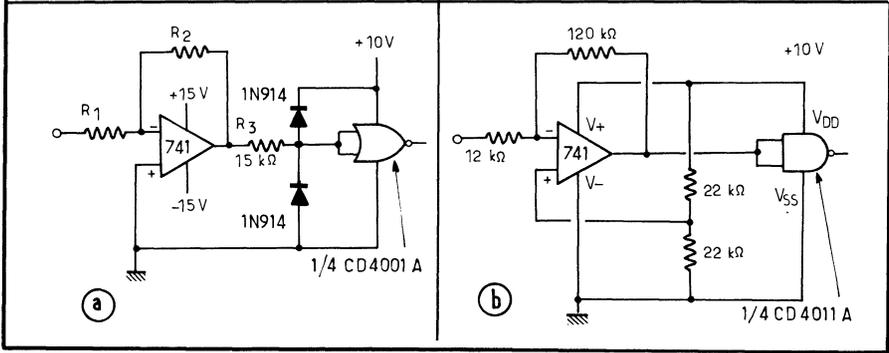


Fig. 22. — Interfaçage amplificateur opérationnel type 741 à CMOS.

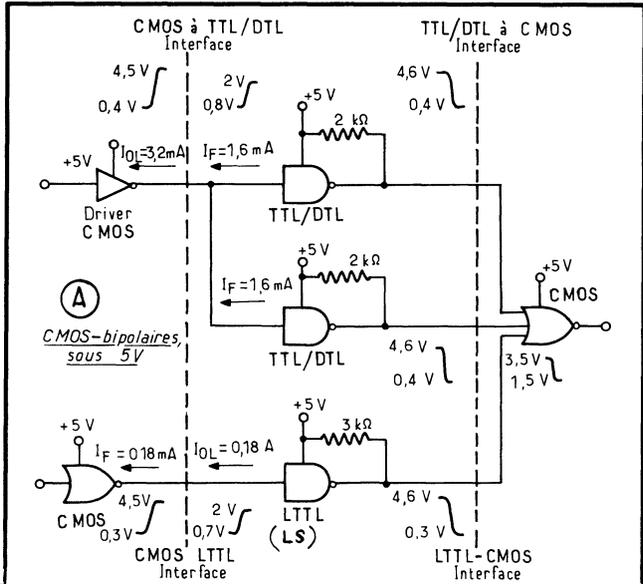
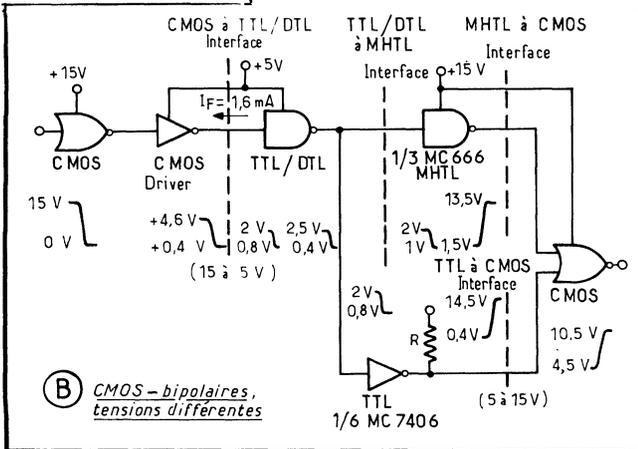


Fig. 23. — Résumé des conditions d'interfaçage CMOS à bipolaires sous des tensions différentes ou identiques (5 V).



sensible (c), mais via un buffer, un thyristor ou triac habituel (d); dans ce dernier schéma, la commande est assurée via un opto-coupleur!

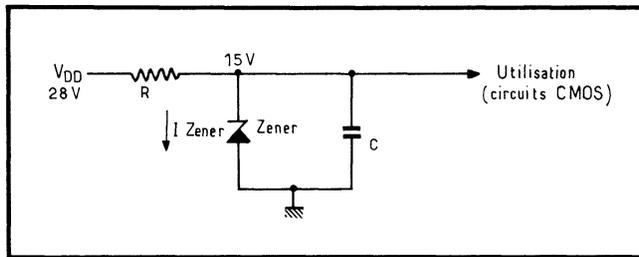
Le CMOS commanderait également via un transistor une diode électroluminescente (LED), mais directement un affichage à cristaux liquides (fig. 21 e, f et g). Enfin, la liaison avec un amplificateur opérationnel, tel le 741, pourrait s'effectuer comme le montrent les schémas 22.

Cette discussion sur les interfaces est résumée, et complétée, par les schémas A à F des figures 23 et 24 (proposés par *Motorola* pour ses MCMOS).

4. LES IMPULSIONS D'HORLOGE

Pour les circuits synchrones, les impulsions d'horloge dont l'excursion doit atteindre la valeur des tensions d'alimentation devront présenter des temps de montée et de chute brefs, n'excédant pas $15 \mu\text{s}$, par exemple, pour les circuits de la série 4000.

Fig. 25. — Principe d'une alimentation ultra-simple pour CMOS.



Ce point est particulièrement important lorsque certains circuits sont connectés en série : flip-flops, registres à décalage... : un temps de transition trop long perturberait le synchronisme et ferait perdre des informations. Il faut, là, que les temps de montée soient inférieurs au temps de propagation auquel s'ajoute le temps de montée du front du signal à la sortie (il dépend de la capacité qui charge le circuit).

5. L'ALIMENTATION DES CMOS

Consommant très peu, les CMOS acceptent également une alimentation non régulée, pourvu que les tensions fournies ne risquent pas d'outrepasser les valeurs limites données par le constructeur (moins de $+3 \text{ V}$ ou au delà de $+15 \text{ V}$, par exemple, pour la série 4000 A).

On notera cependant que la vitesse maximale des circuits s'accroît avec la tension d'alimentation. Cette caractéristique fixera parfois la tension minimale à adopter et, partant, l'excursion possible de la tension d'alimentation.

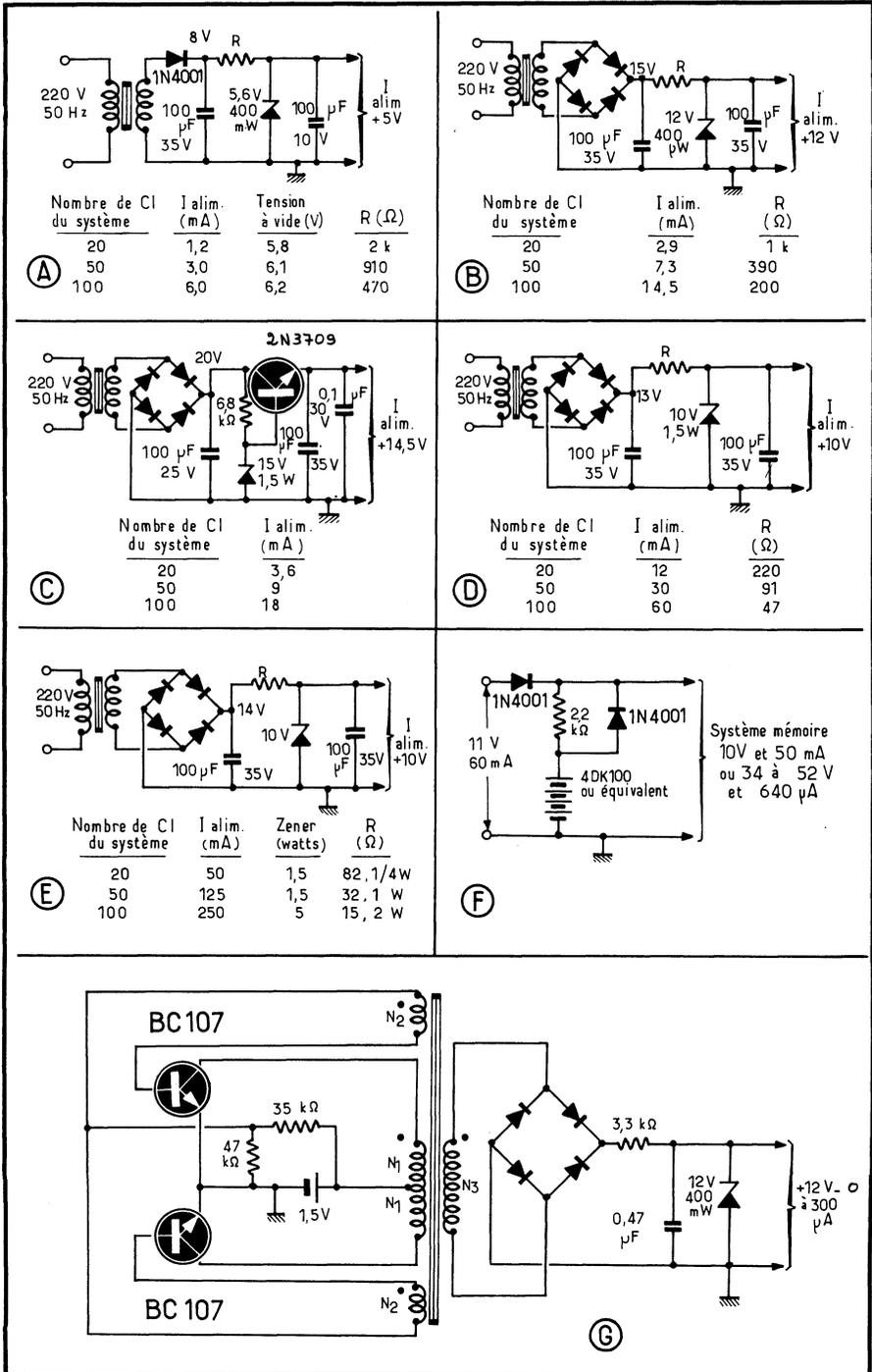


Fig. 26. — Variantes de systèmes d'alimentation, définies dans le texte.

L'utilisation d'une régulation sommaire à diode Zener protégera éventuellement les CMOS des transitoires de ligne, ou autorisera l'utilisation d'une source de tension nominale trop élevée, par exemple le +28 V d'un avion que la régulation à Zener ramènera à +15 V. Le réseau RC devra alors tenir compte des fréquences de travail, les pointes de courant se manifestant lors des commutations (fig. 25). Le filtrage interviendra pour limiter l'ondulation résiduelle. Tant que celle-ci reste inférieure à la valeur de l'immunité au bruit, le circuit CMOS y est insensible; cependant, l'ondulation ne doit pas réduire la tension d'alimentation minimale requise pour une fréquence de travail donnée.

L'utilisation de batteries d'alimentation, même des modèles bon marché, enfin, est rendue plus facile avec les CMOS car ceux-ci supportent de grandes variations de tension. Leur consommation est, on l'a vu, très réduite. De plus, leur immunité au bruit les rend insensibles à l'impédance propre de la source d'alimen-

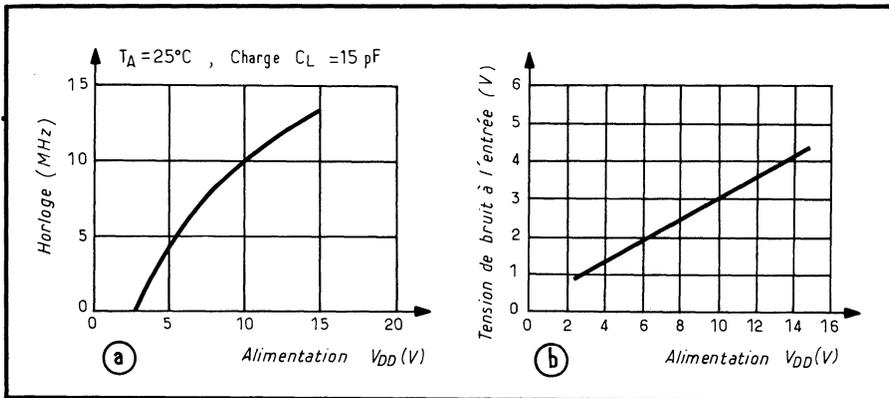


Fig. 27. — L'action de la tension d'alimentation sur la fréquence (a) et le bruit (b).

tation, souvent plus élevée avec des batteries qu'avec des alimentations habituelles. Ainsi, les batteries 12 V d'autos sont directement utilisables, une diode Zener servant éventuellement à éliminer les transitoires accidentelles.

Dans la figure 26 sont dessinées quelques types d'alimentations complètes (d'après D. BLANDFORD et A. BISHOP, note RCA ICAN-6304) avec, par ordre de complexité croissante :

- A. Alimentation pour un système fonctionnant à basse fréquence (moins de 50 kHz) avec une immunité de 1,5 V pour chaque niveau. C'est là un cas d'application très général.
- B. Ici, grâce à une alimentation délivrant 12 V, l'immunité passe à 3,6 V pour chaque niveau.
- C. Avec 14,5 V de sortie, l'immunité est supérieure à 4,35 V garantie : c'est ce que l'on peut obtenir de mieux.

- D. Convient pour un système fonctionnant jusqu'à 2 MHz; en fait, 10 % des circuits devraient travailler à 2 MHz et 90 % à 50 kHz.
- E. En réduisant les valeurs de R, on peut passer à 50 % de circuits à 2 MHz et 50 % à 50 kHz.
- F. Une batterie 12 V sert à l'alimentation, celle-ci n'intervenant qu'en secours en cas de panne du secteur; on préserve ainsi le contenu d'une mémoire RAM de 32 Kbits.
- G. A partir d'une cellule de pile 1,5 V, on recourt à un convertisseur continu-continu. La sortie pourrait attaquer un affichage à cristaux liquides.

Tableau D. — CMOS et capacités internes

CAPACITÉ	SYMBOLE	Valeur ($\mu\text{F}/10^{-3}\text{mm}^2$)	
		CMOS sur silicium	SOS gate silicium
Gate-substrat	C_{GS}	0,4	0,4
Drain-substrat 1 ^(a)	C_{DS1}	0,15	—
Drain-substrat 2 ^(a)	C_{DS2}	0,15	—
Gate-drain	C_{GD}	0,06	0,015

A titre d'exemple, la figure 27 *a* montre comment la fréquence maximale de travail du flip-flop D type CD 4013 A varie avec la tension d'alimentation. En effet, on se souviendra que la fréquence est toujours fonction de l'alimentation. L'immunité, elle, en dépend également et est donnée figure 27 *b* dans le cas le plus défavorable. La dissipation dépend bien entendu aussi de la fréquence (fig. 28), mais aussi de la capacité de charge.

Par sécurité, les bus d'alimentation doivent être découplés par des capacités de 10 nF ou plus, en céramique, toutes les quelque 4 portes, et par 100 nF tous les 4 circuits MSI.

6. LES CMOS/SOS ET LE FONCTIONNEMENT « A FORTE DÉPLÉTION »

Dans le chapitre intitulé « L'arsenal des technologies », on a vu que les CMOS se prêtent à toutes les technologies : gate ou silicium, Locmos, Isoplanar, à isolement diélectrique, SOS... Toutes offrent un intérêt évident, avec une prédominance probablement certaine pour le SOS qui permet un type de fonctionnement original, dit « à forte déplétion ».

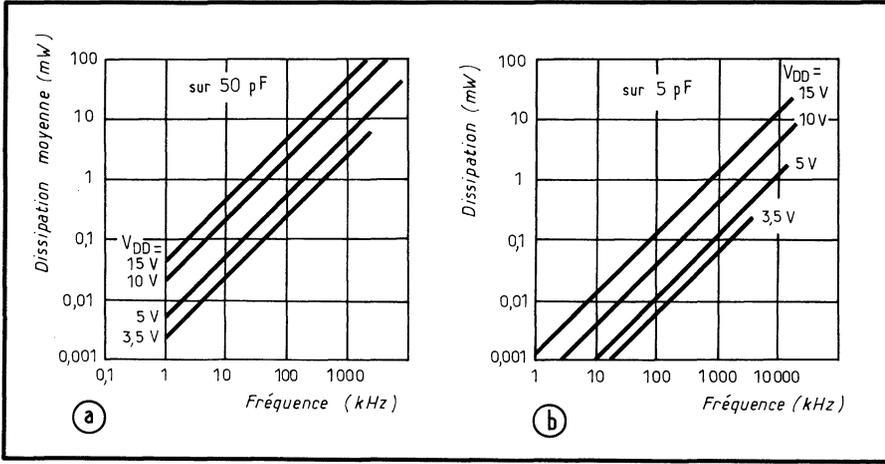


Fig. 28. — Dissipation de la série 4000 en fonction de la fréquence ou de la charge.

Les avantages techniques qu'offrent les CMOS sur saphir (SOS = *silicon on sapphire*) sont indiscutés : à la faible consommation des MOS complémentaires, ils apportent la vitesse, ainsi d'ailleurs qu'une densité d'intégration accrue. Les difficultés résident surtout dans la maîtrise du procédé, et dans le fait que le saphir coûtant plus cher que le silicium, on se demandait longtemps si l'accroissement du prix de vente des CI serait accepté. En 1975, RCA indiquait que le prix de vente des SOS ne dépasserait que de quelque 20 % celui des CMOS sur silicium, différence faible que rend possible un rendement supérieur en production. A l'avenir cette différence pourrait même s'estomper avec le développement de la production et les nouvelles techniques.

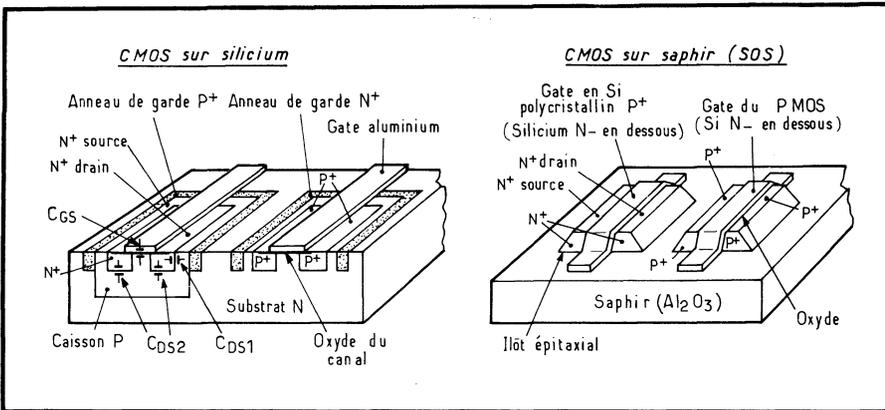


Fig. 29. — CMOS sur saphir (SOS) comparé au CMOS sur silicium.

A propos des SOS, on notera :

- les SOS sont plus rapides parce que des jonctions disparaissent et que les capacités parasites sont considérablement réduites;
- la densité d'intégration est quasiment doublée car on peut supprimer ici les anneaux de garde et rapprocher les MOS; de plus, on dispose de trois niveaux d'interconnexions;
- enfin, les structures sont auto-alignées, ce qui réduit les capacités propres des transistors;
- par contre, le processus exige quelques étapes supplémentaires;
- enfin, la consommation au repos est un peu supérieure à celle des CMOS sur silicium.

La figure 29 compare les deux types de CMOS : on voit que l'avantage de simplicité de structure est acquis pour les SOS. Avec eux, subsiste, en fait, la capacité gate-substrat, ou plutôt gate-canal (tableau D).

Le fonctionnement à forte déplétion

La paire complémentaire de MOS de la figure précédente présente une structure originale que rend possible le SOS. En effet, si le MOS à canal N est classique,

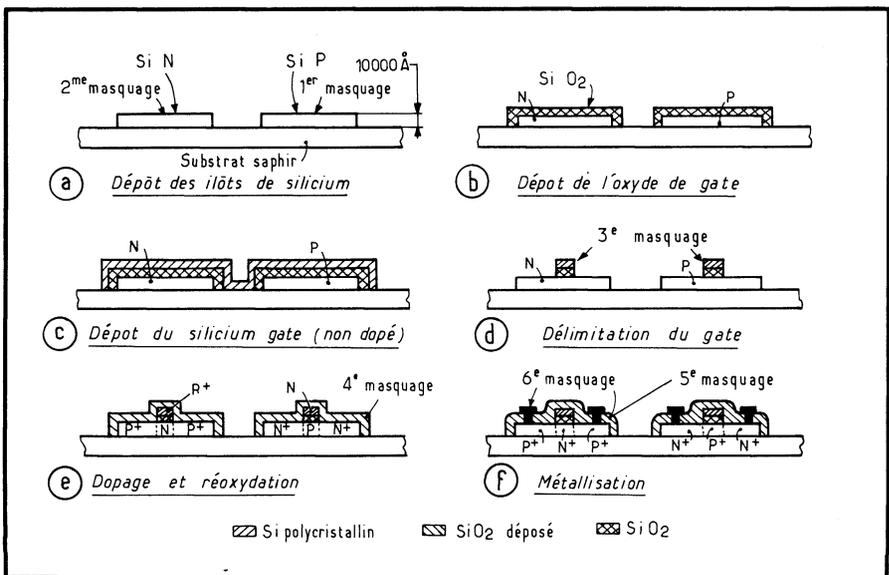


Fig. 30. — Fabrication de l'inverseur CMOS sur saphir.

son voisin ne l'est plus car le canal n'est plus P mais encore une fois N. Pourtant, son fonctionnement obéit aux mêmes lois, car on a jonglé ici avec les niveaux de dopage.

C'est le substrat N^- qu'on utilise habituellement pour le MOS à canal P qui apparaît entre les drain et source N^+ , donc fortement dopés. Avec une tension nulle de gate (condition de blocage), la différence de potentiel de contact entre le gate P^+ et le substrat N^- est de $-0,8$ V et parfaitement suffisante pour chasser les porteurs de la couche épitaxiale et donc l'appauvrir, dans toute son épaisseur. La tension de seuil est alors positive et égale à cette valeur, toute tension plus positive tendant alors à accumuler les charges en surface, donc à rendre le MOS conducteur.

L'avantage qu'offre ce procédé, c'est sa simplicité de réalisation. Pratiquement en effet, le même substrat est déposé en îlots à partir d'une épitaxie pour les deux MOS, et il est également dopé N^- ; par contre et dans la méthode classique, il faut recourir à deux épitaxies différentes N^- et P^- (fig. 30). De plus, on obtient désormais des MOS à faible tension de seuil (entre 0,5 et 0,8 V).

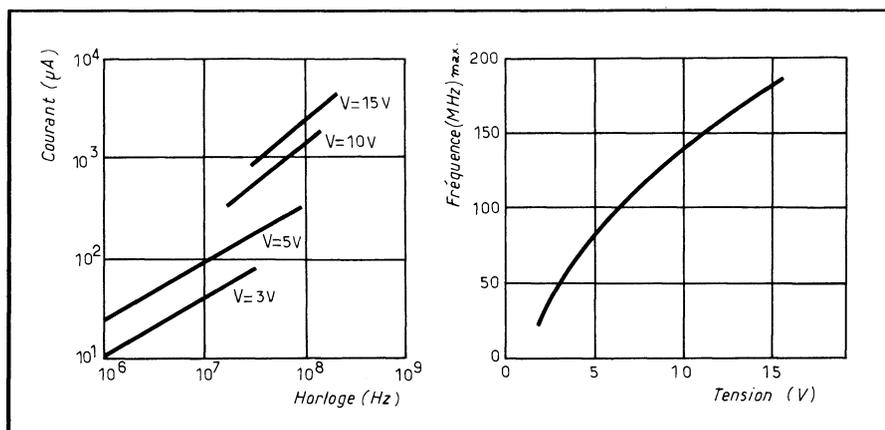


Fig. 31. — Fréquence de travail d'un compteur CMOS en fonction de la tension, et courant consommé.

La figure 31 montre comment varie la fréquence de travail d'un compteur dynamique SOS de ce type en fonction de la tension d'alimentation (en *a*); on atteint 80 MHz sous 5 V et 180 MHz sous 15 V. Le courant s'accroît évidemment avec la fréquence (*b*).

Les niveaux d'interconnexions

Quel que soit le mode de réalisation, à double épitaxie ou à forte déplétion, le SOS dispose de trois niveaux d'interconnexions qui favorisent le LSI :

1. Le niveau dû aux métallisations d'aluminium;
2. Celui offert par le silicium polycristallin servant aux gates;
3. Enfin, le niveau réalisable dans l'épitaxie (forte déplétion), ou les épitaxies.

Ces trois niveaux sont disposés sur le saphir et, de ce fait, aucune capacité parasite n'est à redouter avec le substrat comme lorsque celui-ci est du silicium. Subsiste seule la capacité entre lignes d'interconnexions placées sur un même niveau, mais celle-ci demeure très faible, quoique non négligeable si l'on se préoccupe de vitesse : 0,08 pF environ par millimètre de conducteurs espacés de 7,5 μm (0,002 pF par mil pour espacement de 0,3 mil).

Quelques caractéristiques influant sur la vitesse

La structure des transistors SOS en îlots fait apparaître qu'en fait, chaque MOS se compose de deux éléments en parallèle;

- l'un conventionnel et parallèle à la surface, résultant d'une épitaxie réalisée dans le plan cristallographique $\langle 1.0.0 \rangle$;
- l'autre obtenu sur les flancs de l'îlot et qui se trouve, lui, dans un plan $\langle 1.1.1 \rangle$.

Or, la charge Q_{SS} fixée dans l'oxyde créé par croissance thermique est supérieure à celle provenant d'un oxyde obtenu par oxydation du silicium. En conséquence, une plus faible tension appauvrit le substrat épitaxial et la tension de seuil décroît sur les flancs du silicium N, de 0,3 V environ, par rapport au seuil en surface. L'inverse est vrai pour le MOS à canal P.

En moyenne, la tension de seuil est faible (0,5 à 0,8 V), ce qui est ainsi favorable à la vitesse. Un autre élément important jouant sur la vitesse est l'intensité que peut débiter un transistor, celle-ci dépendant de la mobilité des porteurs. Or, avec le SOS, le courant est très proche sinon égal à celui qu'on obtient de MOS sur silicium.

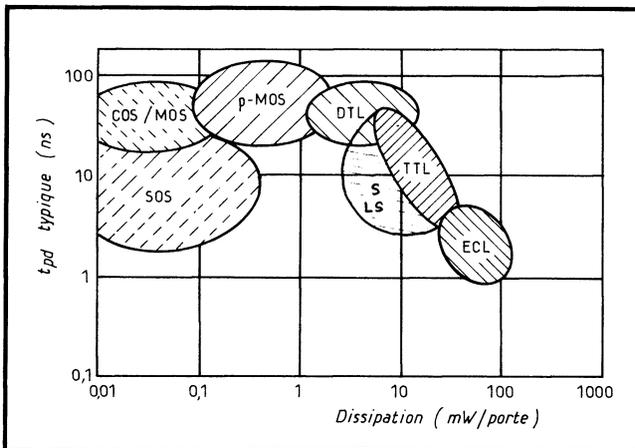


Fig. 32. — Dissipation comparée de diverses logiques.

Enfin, la réduction des capacités internes est telle que sur une chaîne de portes NOR intégrées, des temps de propagation de 1 à 2 ns peuvent être obtenus sous 10 V, avec des temps de chute et de montée de 4 ns. Ces résultats sont comparables à ceux qu'on relève avec d'excellentes logiques Schottky. En pratique, cependant, ce sont fréquemment les capacités externes (dues aux connexions) qui limitent la vitesse de travail.

Comme on vient de le voir, un compteur SOS dynamique peut atteindre 80 MHz sous 5 V, toute en ne consommant alors que 2 mW. La TTL standard, 54/74, atteint, elle, quelque 30 MHz sous 5 V, pour une consommation de 200 mW, avec une immunité de l'ordre du volt, contre plus de 2 V pour les CMOS.

La figure 32 compare les facteurs de qualités des SOS et d'autres familles logiques classiques (d'après S. Sheffield EATON, de RCA Somerville, USA). Quant à la surface occupée, un inverseur MOS sur saphir se contente de quelque $1\,250\ \mu\text{m}^2$ (2 mils²) contre $2\,500\ \mu\text{m}^2$ (4 mils²) pour un inverseur MOS sur silicium, avec tous les contacts de sorties.

La variante ESFI, de Siemens

Pour accroître la densité d'intégration comme pour diminuer les courants de fuite drain-source des MOS sur saphir qui, en 1976, restent sensiblement supérieurs à ceux des MOS sur silicium, la société *Siemens* a proposé une structure d'inverseur à un MOS, plus sa résistance de charge. La cellule statique qui en résulte est dessinée figure 33; on la comparera éventuellement avec les schémas de cellules-mémoires données dans le chapitre consacré aux mémoires.

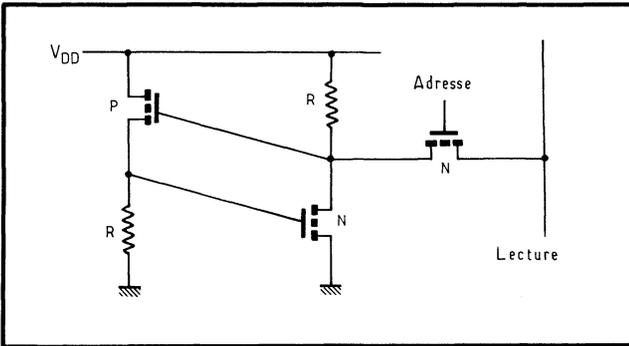


Fig. 33. — La porte ESFI (sur saphir) proposée par Siemens.

Une telle cellule réduit les problèmes d'interconnexion en supprimant les croisements, ce qui est important en SOS (dont ESFI est la traduction allemande : *Epitaxial Silicon Film on Insulator...*). Le transistor MOS d'adressage pourrait même être remplacé ultérieurement par une diode.

Tableau E. — CMOS et SOS comparés aux familles classiques

FAMILLE	t _{pd} (ns)	F	CONSUMMATION	IMMUNITÉ	SORTANCE
		FLIP-FLOP (MHz)	AU REPOS (mW)	AU BRUIT (V)	
TTL standard 54/74	10	35	10	1	10
54L/74L	33	3	1	1	10
DTL	30	5	8	1	8
Schottky •74 LS	10	40	2	0,8	20
•9LS(Fairchild)	5	80	2	0,8	20
CMOS 4000A • 5 V	35	5	10 nW	2	50
	• 10 V	25	10 nW	4	50
CMOS/SOS • 5 V	10	50	15 nW	2	50
	• 10V	5	15 nW	4	50
Schottky standard 54S/74S	3	135	20	0,8	10
ECL 10 K	2	200	25		

Tableau F. — Les SOS « durcis » comparés à d'autres circuits.

CARACTÉRISTIQUES	TTL A ISOLEMENT DIÉLECTRIQUE DURCIE	PMOS STANDARD	PMOS SOS DURCIS	PMOS SOS DURCIS
Dose de neutrons tolérée (rads)	3×10^{14}	10^{13}	5×10^{15}	5×10^{15}
Ionisation tolérée (rads)	$>10^7$	10^4	$>10^7$	$>10^7$
Vitesse	25MHz	2,5MHz	20MHz	60MHz
Puissance consommée	10mW	0,5mW	0,5mW	0,2mW

La résistance aux radiations ionisantes

Un aspect quelque peu particulier des circuits intégrés réside dans leur sensibilité aux rayonnements ionisants. Ceux-ci, en effet, créent des paires électron-trou dans le silicium et perturbent ainsi le fonctionnement des circuits, ou effacent les mémoires. Ce risque est considérablement diminué si le substrat est du saphir, isolant.

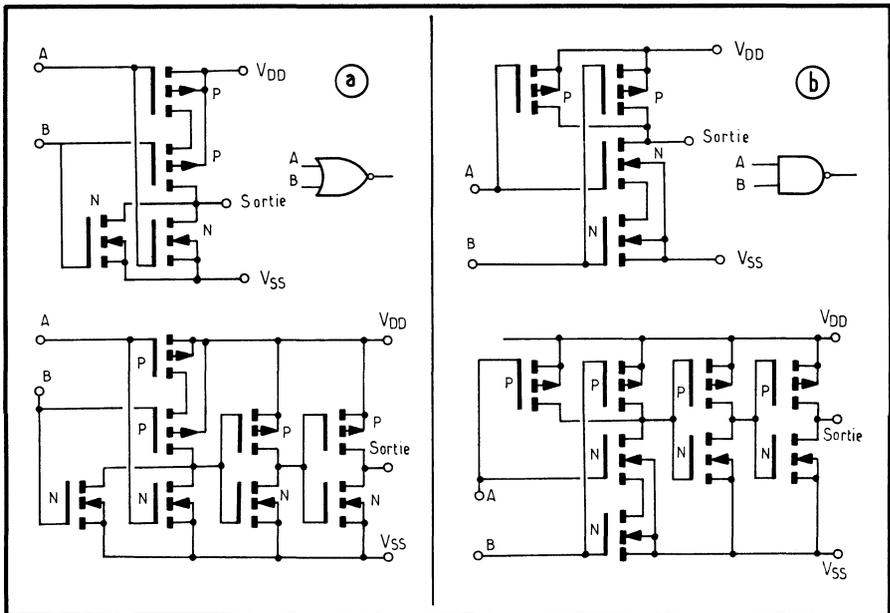


Fig. 34. — Sorties CMOS sans (a) et avec (b) buffers.

De ce fait, les SOS dont on a accru encore la résistance aux rayonnements par des procédés complémentaires — on dit qu'ils ont été « durcis » — s'avèrent comparables aux meilleures TTL à isolement diélectrique et « durcis » (tableau F, selon *Rockwell*).

Les circuits « durcis » s'appliquent essentiellement au domaine militaire, à l'aérospatial, et aux applications visant le nucléaire.

7. LA 54 C/74 C DE NS

La série CD 4000 de RCA sert de référence en MOS complémentaires. Mais d'autres familles, variantes ou non, et recourant à toutes les technologies existantes : les séries *Motorola* (14000 et gate au silicium), *Fairchild* (Isoplanar 34000), *Harris Semiconductor* (à isolement diélectrique DI/CMOS), *Locmos de RTC* et *Philips*, ou encore type 54C/74C de *National Semiconductor* (NS)...

Cette dernière société a en effet introduit une famille CMOS compatible 54/74, c'est-à-dire qu'elle a réalisé des circuits à fonctions, brochages et tensions identiques, réplique exacte en CMOS de ce qui existe en 54/74 TTL.

L'objectif est de faciliter le passage des utilisateurs à la CMOS sans reversion des schémas, mais avec le bénéfice immédiat résultant de l'utilisation des CMOS.

8. LA 34000 DE FAIRCHILD ET LA LOCMOS DE RTC

On vient de citer, parmi les variantes technologiques, la 34000 Isoplanar de *Fairchild*. Cette famille illustre parfaitement les avantages des technologies évoluées, et l'on a vu dans le chapitre consacré à l'arsenal des technologies qu'elle permettait un gain de surface de 35 % sur les structures CMOS à gate aluminium.

Ici, tous les circuits sont de plus et systématiquement dotés de buffers de sortie, donc de deux inverseurs en série (fig. 34) qui accroissent la puissance de sortie et par conséquent, la vitesse, comme le montre la figure 35 a en fonction de la

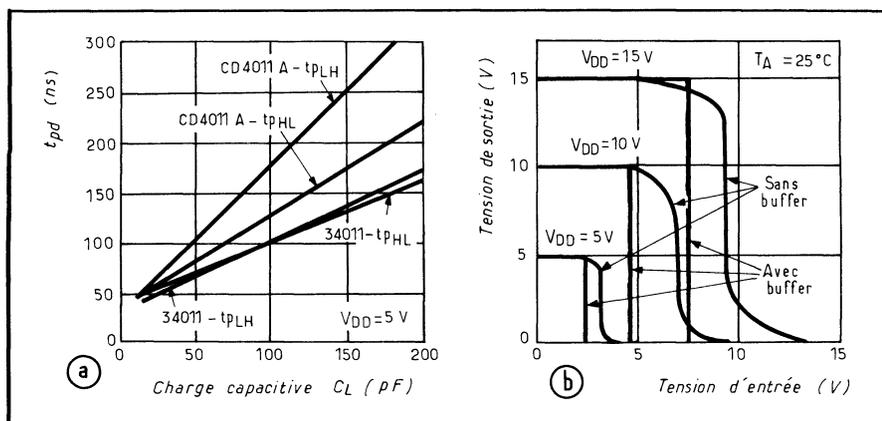


Fig. 35. — Courbes typiques relatives à des portes NAND dotées de buffers de sortie, comparées aux portes identiques de la série 4000 sans buffers.

charge capacitive : la différence est sensible pour de fortes charges mais s'estompe totalement vers les 10 pF. En *b*, on voit par contre que les transitions des fonctions de transfert sont plus rectangulaires. La LOCMOS de *RTC* applique ces mêmes principes.

9. L'ISOLEMENT DIÉLECTRIQUE

Ainsi qu'on l'a vu dans le chapitre consacré à l'arsenal des technologies, l'isolement des CMOS fondé sur le principe des jonctions polarisées en inverse peut conduire certaines structures à se manifester comme des thyristors.

Le même phénomène peut être dû aux diodes de protection des entrées (fig. 36). Une transitoire importante de bruit à l'entrée, se traduisant par une tension inférieure à la masse ou supérieure à V_{CC} , polarisera l'une des diodes dans le sens passant; cette diode et la jonction PN du substrat se comportent alors comme un semiconducteur à 4 couches, c'est-à-dire un thyristor qui entre en avalanche; le courant I_{CC} s'accroît jusqu'à devenir destructif si aucune limitation n'intervient.

La protection peut être alors assurée :

- soit en limitant le courant d'entrée, à l'aide d'une résistance; c'est la solution la plus communément employée;

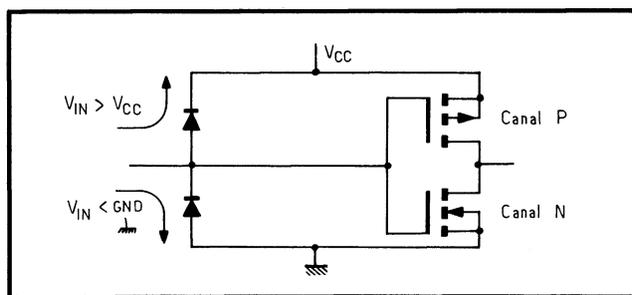


Fig. 36. — Les diodes d'entrée peuvent être à l'origine de blocages intempestifs.

- soit en limitant rigoureusement les tensions d'entrée;
- soit en recourant à un isolement autre que par jonctions polarisées en inverse.

Le SOS répond parfaitement à cette dernière exigence; l'isolement diélectrique, à l'aide de silice, également. Cette dernière formule, qu'il convenait de rappeler ici, a été développée dans le même chapitre consacré aux technologies. La firme *Harris Semiconductor* s'en est fait le champion.

10. LE SUBSTRAT EN SPINELLE

N'ayant pas encore donné lieu jusqu'en 1978 à des réalisations commerciales, les CMOS sur substrat en spinelle, appelés par conséquent également SOS, paraissent eux aussi forts intéressants. En effet :

1. Le réseau du spinelle est encore plus proche du silicium que celui du saphir.

2. Le spinelle à la transparence du verre et pourrait servir en optoélectronique intégrée.
3. Le spinelle est plus facile à usiner.
4. Enfin, le spinelle pourrait également se prêter à des réalisations vers les micro-ondes.

11. CMOS OU I²L?

En fait, le concurrent essentiel des CMOS pourrait bien être la logique bipolaire à injection I²L (de *integrated injection logic*), commercialisée pour la pre-

Tableau G. — I²L, CMOS et TTL

PARAMÈTRES	I ² L (1975)	CMOS (sous 5V)	TTL (standard)
Densité d'intégration (portes/mm ²)	120 à 200	70 à 160	20
Tensions d'alimentation (V)	0,85 à ∞ (*)	1,5 à 16	5
Temps de propagation (ns)	25 à 250	30	10
Fréquence maximale (MHz)	5	8	30
Consommation statique	1 nA à 1 mA 1 nW à 5 mW	2 nA 10 nW	2 mA 10 mW
Consommation dynamique (5 MHz)	1 nW à 5 mW	0,1 mW	10 mW
Facteur de qualité (P ^d)	0,1 à 4	1 à 5	100
Excursion logique (V)	0,2	5	3,5
Immunité au bruit (V)	0,03 à 0,3	2	1
Sortance avec même famille	5	50	10
Nombre de masquages	5	6	7
Nombre d'entrées par porte	1	>1	>1
Nombre de niveaux d'interconnexions	1	3	2

(*)_Une résistance chutrice externe commande le courant injecté.

mière fois en 1975. Dans ses premières manifestations, l'I²L fournit la même densité d'intégration que les CMOS mais une vitesse moindre pour une consommation un peu supérieure. La figure 37 présente ses traits essentiels, et le tableau G compare ses performances à celles des CMOS et de la TTL 54/74 standard, pour mieux situer les grandeurs; on notera, à ce propos, que les valeurs données ne devront jamais être considérées comme absolues (ce qui explique les écarts que l'on pourra constater d'un tableau à l'autre, en fonction de la comparaison effectuée).

La figure 38 compare les surfaces de silicium occupées par une porte élémentaire, et ce dans plusieurs familles avec, ici, la I²L dans sa version « non isolée » (sans diode d'isolement). Son avantage n'est cependant pas évident, ici, car le problème des interconnexions cité dans le tableau, pourrait bien le remettre en cause.

12. LOGIQUE MAJORITAIRE ET CMOS

C'est en familles CMOS que la logique majoritaire s'est d'abord manifestée sous forme de circuits intégrés; c'est pourquoi sa présentation a été incluse dans ce chapitre.

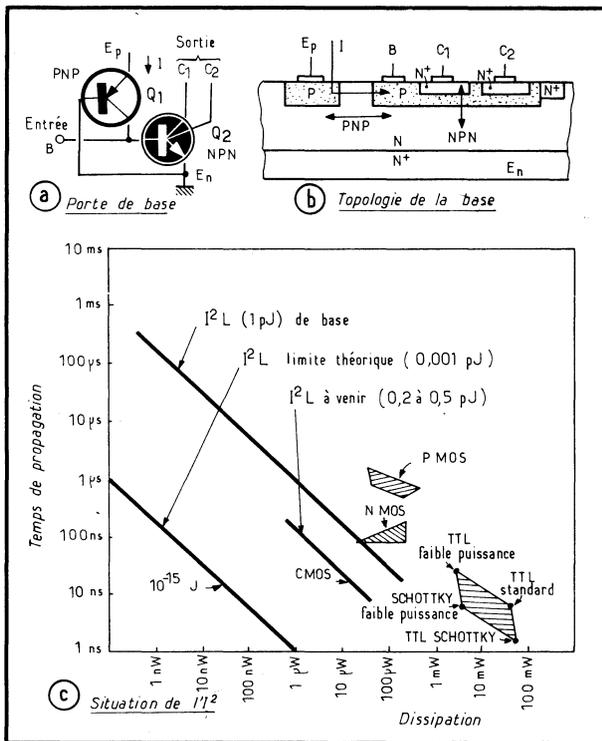


Fig. 37. — Quelques notions sur la I²L.

Comme son nom l'indique, la logique majoritaire dégage une condition « vraie » lorsque la majorité des signaux d'entrée est au niveau 1. Favorisée par le développement des CMOS, cette forme de logique de « vote » pourrait s'appliquer à la corrélation entre mots, aux comparaisons pondérées et vérifications de parité, à l'identification des caractères (en lecture optique, par exemple), à l'amélioration des échos radar, la reconnaissance de codes, l'extraction des signaux du bruit, et à tout traitement de caractère probabilistique.

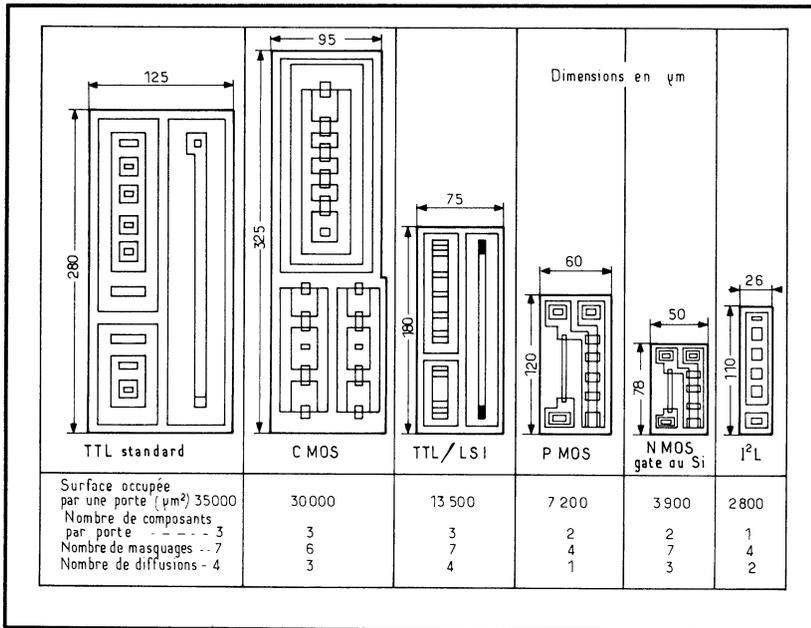


Fig. 38. — Topologie de quelques cellules de base pour diverses familles.

Ainsi, elle a pu être exploitée en Allemagne, par *AEG-Telefunken*, en logique de décision pour assurer la sécurité des chemins de fer : trois signaux qui devraient, en principe, être identiques si aucune défaillance d'équipement ne s'est produite, déclenchent une décision majoritaire. La défaillance d'un seul équipement sera sans effet et pourra être signalée pour réparation.

Principe de la logique majoritaire

En logique courante, les fonctions de base sont les NOR et NAND, dérivées des OR et AND, qui fondamentalement détectent une combinaison logique des entrées et une seule. Il suffit de se reporter aux tables d'états bien connues de ces logiques pour le vérifier. Ainsi, chaque porte détecte 1 combinaison sur 2^n où n est le nombre des entrées.

Avec les fonctions exclusives (OR et NOR exclusifs), la reconnaissance porte sur n combinaisons sur les 2^n possibles, n étant toujours le nombre des entrées.

La logique majoritaire, elle, détecte une situation majoritaire, c'est-à-dire, en toute bonne logique, 2^{n-1} combinaisons sur les 2^n possibles. Il est évident que si $n = 2$, la majorité s'exprimera lors de 2 combinaisons d'entrée sur les 4 possibles; si $n = 3$, elle portera sur 4 combinaisons sur les 8 possibles, etc.

Cette forme de logique n'avait pas débouché jusqu'à présent en raison de la complexité de sa réalisation. Il en est allé de même avec bien d'autres types de circuits, tels que boucle à phase asservie, mémoires associatives, etc. Prévoir l'ensemble de ses applications possibles semble alors encore prématuré.

Pourtant, le développement des familles CMOS a permis de redonner un souffle nouveau à la logique majoritaire, dont *Motorola* semble vouloir se faire

le héraut. Cette firme proposait le MC 14530, double porte à logique majoritaire à 5 entrées. Ajoutons que cette logique majoritaire s'inscrit dans ce que l'on désigne par « logique à seuil » pour laquelle RCA proposait, dès 1967, une porte en composants discrets (décrite par R. Damaye dans « Logique électronique et circuits intégrés numériques », volume édité par la Société des Éditions Radio).

La porte MC 14530

Le circuit intégré MC 14530 constitue le type même de la porte en logique majoritaire, aussi allons-nous nous y référer. Ce circuit est une double porte à 5 entrées (fig. 39); outre les portes majoritaires notées M_5 , apparaissent des NOR exclusifs dont chacun, grâce à l'entrée de commande W :

1. Inverse la fonction selon que $W = 1$ ou $W = 0$. Donc, on détecte une situation minoritaire par une sortie vraie, pour $W = 0$.

2. Transforme éventuellement la porte en OR, AND, NOR ou NAND.

C'est ce que montre la table d'états jointe à la figure, qu'en tout état de cause la fonction principale reste la fonction majoritaire qui se traduit par une sortie = 1 lorsque 3 des entrées sur 5 sont à 1 (à la condition que $W = 1$).

Électriquement, une telle porte se traduit en CMOS par le schéma de la figure 40 qui témoigne bien de sa complexité relative. Les 6 « sections » complémentaires réalisent toutes les combinaisons possibles; dès que 3 entrées sont au niveau 1, une branche verticale reliée à la masse est débloquée et porte la sortie du « groupe » M_5 à 0, ce que l'on peut vérifier facilement. À l'inverse, s'il n'y a

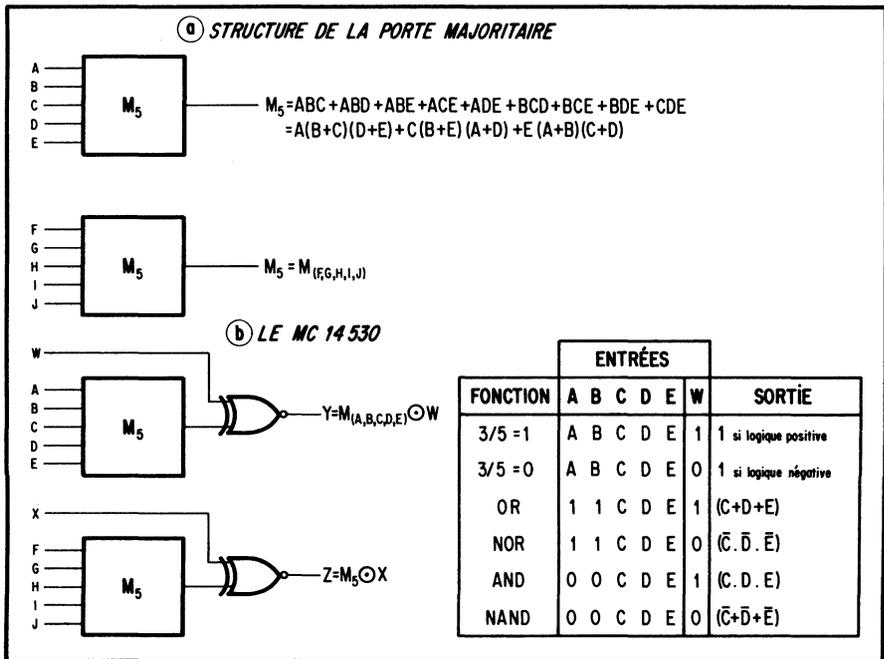


Fig. 39. — Symbole de la porte majoritaire (a) et structure du circuit intégré MC 14530 (b).

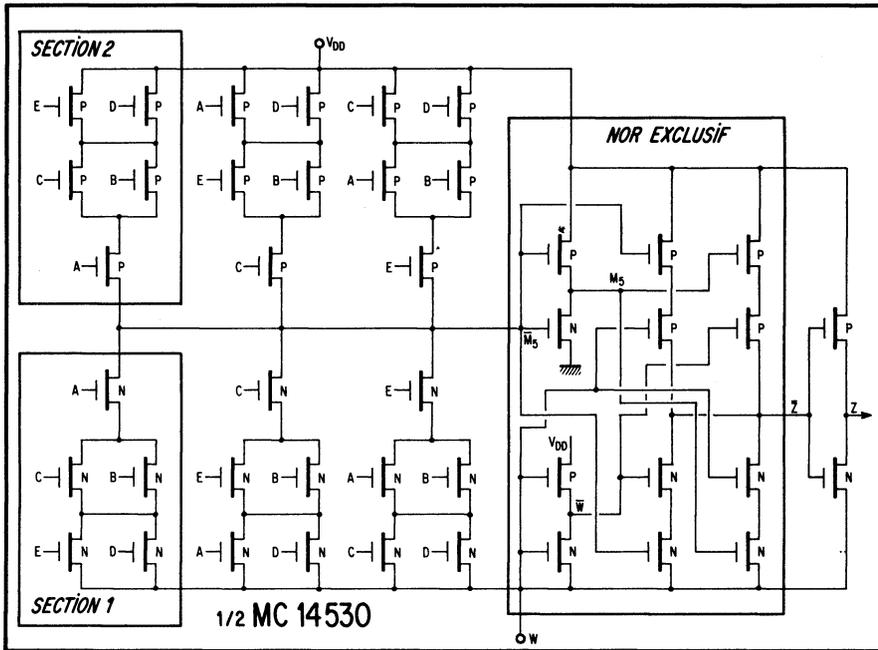


Fig. 40. — Schéma électrique du MC 14530 (un demi), porte majoritaire à 5 entrées et entrée de commande.

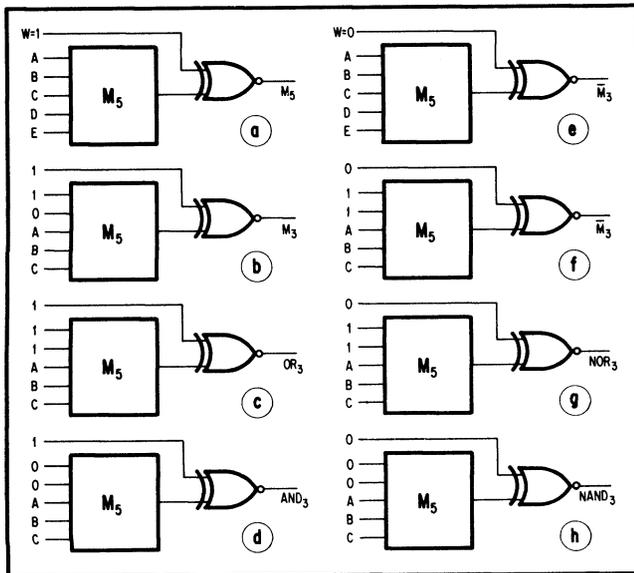


Fig. 41. — Montages de base du MC 14530 : décisions majoritaires ou minoritaires) sur 5 ou 3 entrées; réalisation des fonctions logiques habituelles.

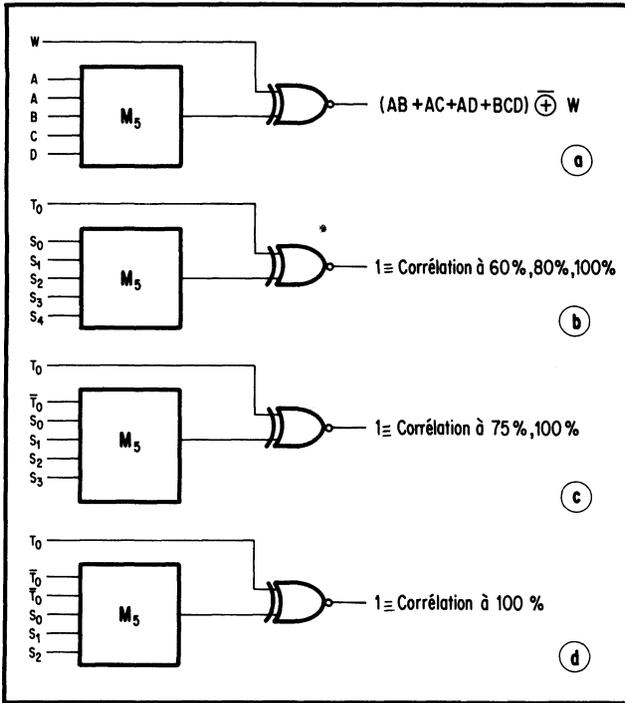


Fig. 42. — Comment on double le poids relatif d'une entrée (a). Applications à la corrélation (b à d).

Fig. 43. — Flip-flops et multivibrateur dont la décision dépend ou non d'une coïncidence.

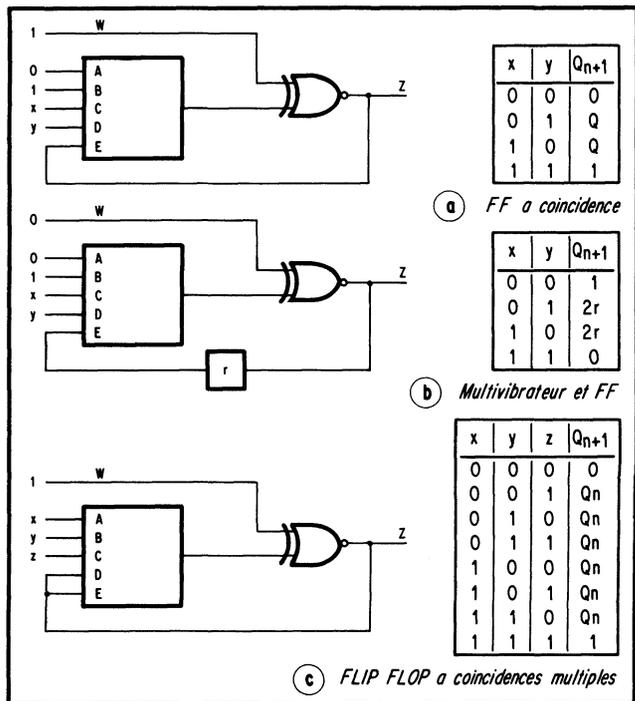


Fig. 44. — Exemple de fonction complexe réalisé à l'aide d'un seul MC 14530.

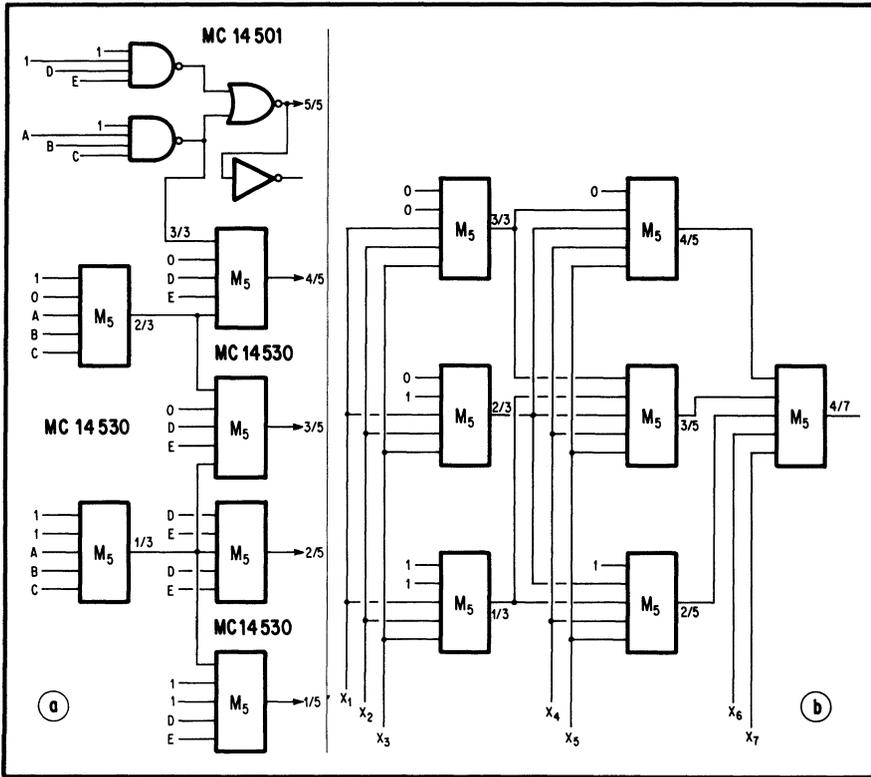
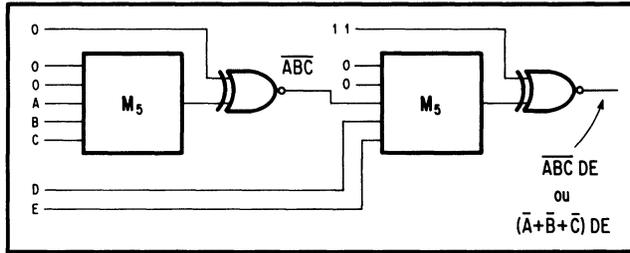


Fig. 45. — Deux exemples de réseaux à logiques majoritaires pouvant détecter le nombre d'entrées vraies (ou 1).

pas majorité de 1, on trouve plus de 3 entrées à 0 qui débloquent une chaîne de MOS canal P et portent donc la sortie de M_5 à 1. On raisonne, ici, en termes de logique positive; la sortie de M_5 est l'inverse et doit donc être rétablie par un inverseur.

Le NOR exclusif qui suit ne comporte que 12 transistors MOS, y compris deux inverseurs, et est suivi par l'inverseur de sortie.

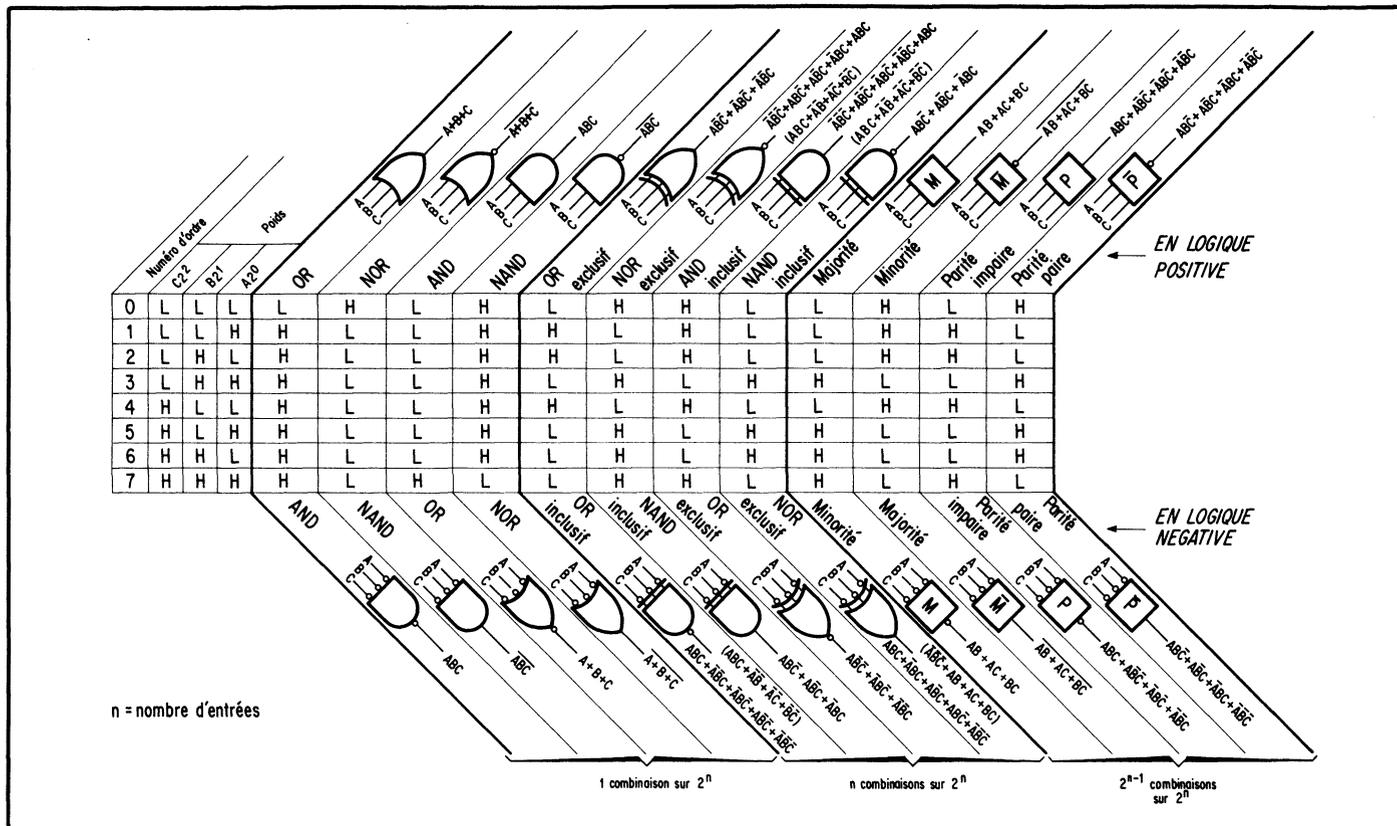


Fig. 46. — Tableau récapitulatif des fonctions de logique majoritaire, sur 3 entrées.

Sous 10 V, une telle porte offre un temps de propagation de 100 à 200 ns selon combinaison, et ne consomme que quelque 25 nW au repos. Son impédance d'entrée est celle des CMOS : environ $10^{12} \Omega$, et sa sortance est supérieure à 50. Enfin, l'immunité au bruit est habituelle, 45 % de V_{DD} .

Les applications

Les applications que l'on va examiner ici montrent bien la puissance de la porte à logique majoritaire. Ainsi, on trouve figure 41 :

- En (a), la détection d'une majorité sur 5 entrées.
- En (b), la détection d'une majorité sur 3 entrées seulement.
- En (c), un OU avec 3 entrées.
- En (d), un AND sur 3 entrées.

De (e) à (h), des conditions inverses car on a fait $W = 0$, donc :

- En (e), un état minoritaire sur 5 entrées (traduit par une sortie à 1).
- En (f), un état minoritaire sur 3 entrées.
- En (g), un NOR sur 3 entrées.
- En (h), un NAND sur 3 entrées.

Si l'on désire accroître le « poids » d'une variable, il suffit de l'appliquer à 2 entrées (pour doubler son importance), comme le montre la figure 42 a.

La corrélation d'un bit de test noté T_0 (en b) s'effectue à 60 %, 80 % ou 100 % selon qu'elle s'effectue avec 3, 4 ou 5 entrées; en (c), elle sera de 75 % ou 100 % puisque l'inverse du bit de test est appliqué à l'une des entrées de M5. Pour une raison semblable, elle ne peut plus être que de 100 % en (d), toujours pour une sortie au niveau 1.

A l'aide de tels montages, on corrèle des mots ou des données, même entachées de bruit. La lecture optique ou la reconnaissance de forme, l'amélioration d'échos radar ou autres peuvent en constituer des applications.

La figure 43 montre, en (a), un flip-flop qui ne change d'état que lorsque les deux entrées changent simultanément d'état (coïncidence). En (b), cette bascule est transformée en multivibrateur (astable) lorsque les entrées X et Y ne sont pas en coïncidence; la période est le double du tpd de la porte auquel s'ajoute le retard dans la ligne de réaction Z vers V. En (c), le flip-flop ne change d'état que pour une situation « tous zéros » ou « tous uns »; on dispose ainsi d'un moyen facile de détecter de telles situations.

La figure 44 illustre une fonction complexe, résultant d'un NAND sur trois variables, suivi par un AND avec deux autres variables. On pourrait multiplier à l'infini de tels exemples.

La porte majoritaire peut servir également à détecter si n bits sur les m d'un « mot » donné sont au niveau 1; la figure 45 montre ainsi comment en détecter 1/3, 2/3, 3/3 ou de 1/5 à 5/5, ou même 4/7 bits au niveau vrai.

Ce rapport n'est pas limitatif, on s'en doute; un schéma plus général et correspondant à tous les cas possibles de n/m demandera, pour une détection précise de faire suivre ce montage par un codeur prioritaire qui fournira en sortie le nombre exact d'entrées (n) au niveau 1, ce nombre étant codé sur 3 bits.

Les techniques de corrélation mènent à des schémas plus complexes, mais parfaitement réalisables. La figure 46 enfin, résume les états de logiques majoritaires relatifs à des fonctions de base, et ce pour 3 entrées.

TROISIÈME PARTIE

LES APPLICATIONS

CHAPITRE XIV

LES MÉMOIRES MORTES

Les mémoires constituent une importante application des circuits intégrés à transistors MOS. On va les passer en revue en commençant par les mémoires mortes. Ce sont des mémoires qui, en service normal, ne peuvent être que lues. Leur type de base est la ROM, ou mémoire à lecture seule (read only memory); elle est préparée et programmée par le fabricant de circuits intégrés, par grandes séries. Sa variante PROM (ROM programmable) est, elle, programmable à l'unité par l'utilisateur. Enfin, la EPROM (PROM programmable électriquement), appellation synonyme en pratique de REEPROM (PROM reprogrammable), est une PROM qui peut être éventuellement effacée et reprogrammée par l'utilisateur; ses variantes « tout électrique » sont les EAROM et EEROM. On examinera encore ici le cas des réseaux logiques reprogrammables, ou PLA.

1. LE CLASSEMENT DES MÉMOIRES

Le chapitre des mémoires est extrêmement vaste et complexe. Par souci de simplification, on peut ranger les mémoires dans six grandes catégories :

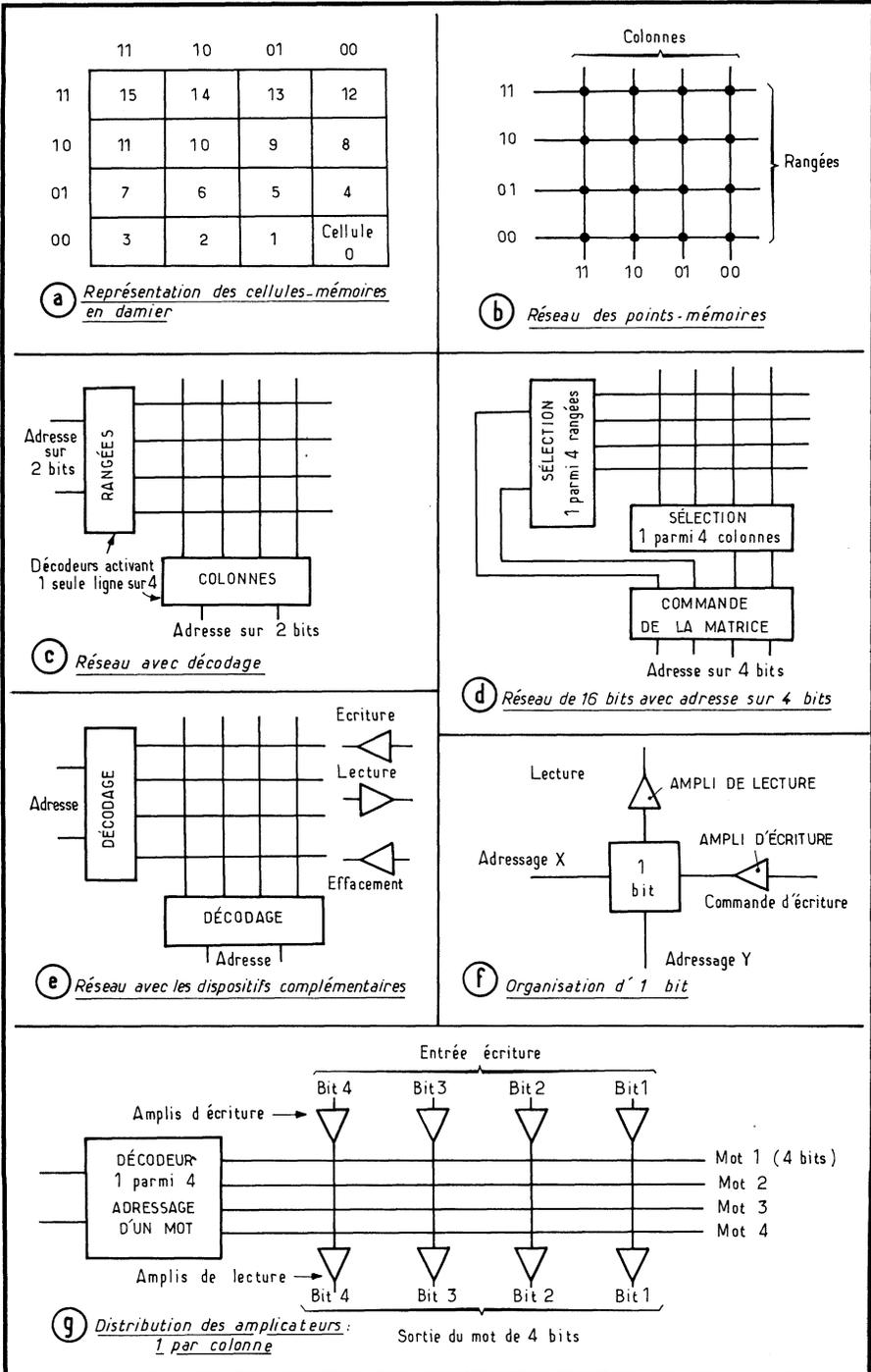


Fig. 1. — Structure d'une mémoire.

1. Les RAM, comparables dans leurs fonctions aux tores de ferrites, donc à lecture, écriture et effacement. Elles sont adressables, c'est-à-dire qu'on peut désigner précisément n'importe quelle cellule élémentaire de la mémoire par son « adresse » et agir exclusivement sur elle; on dit de façon courante (mais cette formulation est discutable) que la mémoire est à *accès aléatoire*, ce qui traduit RAM : *random access memory*.

2. Les ROM, adressables elles aussi mais à lecture uniquement. Elles sont programmées, c'est-à-dire enregistrées, ou inscrites, par le fabricant de circuits intégrés qui ne peut procéder que par de grandes séries identiques pour que la mémoire soit économiquement viable. Le sigle ROM provient de *read-only-memory* : à lecture uniquement. On l'appelle également « mémoire morte ».

3. Les PROM : il s'agit de ROM qui, cette fois, sont enregistrées à l'unité par l'utilisateur, et donc livrées vierges par le fabricant. Le sigle PROM provient de *ROM Programmable*.

4. Les REPROM : ce sont des PROM que l'on peut occasionnellement effacer (mais occasionnellement cependant car cette opération peut se révéler longue ou peu commode), pour ensuite les réinscrire, d'où leur sigle.

5. Les registres à décalage : ici, l'information passe d'une cellule de registre à la suivante. Fondamentalement, on lit l'information « au passage », comme une tête de lecture lirait une piste de disque magnétique. L'accès est donc séquentiel et non plus aléatoire.

6. Les mémoires adressables par le contenu (CAM); leur accès n'est donc plus aléatoire; elles sont surtout utilisées en mémoire-carnet dans les unités centrales d'ordinateurs.

2. L'ORGANISATION D'UNE MÉMOIRE

La mémoire stocke des informations binaires, 0 ou 1; l'unité est donc le bit :

- *Le bit* (de « binary digit ») est l'unité d'information numérique binaire.
- *Le digit* (digital unit) est l'unité générale d'information numérique.
- *Un mot* se compose de un ou plusieurs bits; leur nombre doit être précisé dans chaque cas. Ainsi, on trouvera des mots de 1 bit, 2 ou 4 bits, 8, 16 ou 32 bits... et toutes autres valeurs intermédiaires ou supérieures.
- *Un octet* est un mot de 8 bits.
- *Le « byte »* américain qui traduit le *mot* se voit souvent attribuer la valeur d'un octet.

• *Un kilo*, préfixe qui vaut 1 000 en décimal, et est représenté par un *k* minuscule, vaut 1 024 en binaire pour suivre la progression logique des valeurs. On le représentera ici par un *K* majuscule. Ainsi, une mémoire de 4 K bits est une mémoire de 4 096 bits; 2 K octets représentent 2 048 mots de 8 bits...

La mémoire intégrée est constituée par un réseau de points-mémoires, des cellules élémentaires qui emmagasinent des 0 ou des 1. On peut la représenter par le damier de la figure 1 *a*, pour une mémoire de 16 bits en réseau de 4×4 . La structure correspondante est donnée en *b*, où chaque point de croisement est un point-mémoire. L'adresse de chaque point est ainsi précisée en binaire pour les rangées et les colonnes, le point visé étant à leur intersection (chaque point-mémoire contenant un 0 ou un 1).

L'adressage horizontal et vertical peuvent donc être assurés, dans le cas d'une mémoire 16 bits, par deux mots de 2 bits, pourvu que ceux-ci actionnent des

décodeurs à 1 parmi 4 qui activeront chacun une seule ligne sur les 4 qu'ils pilotent, et cela, en fonction du mot reçu (*c*).

L'adresse d'entrée peut donc être regroupée sur 4 bits (*d*), l'utilisateur n'ayant pas à se soucier, à la limite, de l'organisation interne du réseau-mémoire. Par exemple, la RAM type TMS 4030 est une mémoire de 4 096 mots de 1 bit seulement, mais son réseau interne comprend deux matrices de 32×64 cellules. Il est évident que l'adressage des 4 096 cellules exige un mot de 12 bits (pour répondre aux 4 096 combinaisons : $2^{12} = 4\,096$).

En (*e*), enfin, on a symbolisé les circuits qui servent, selon les cas, à l'écriture, la lecture ou l'effacement. Ainsi, chaque point-mémoire est attaqué, en principe, par 4 lignes pour une RAM (*f*), l'effacement pouvant consister en l'inscription d'un zéro. Chaque amplificateur dessert cependant toute une rangée, ou toute une colonne (*g*); en principe, on trouvera autant d'amplificateurs d'écriture ou de lecture que le mot comporte de bits.

Dans le cas de la figure, la mémoire ne comprenant que 4 mots de 4 bits, le choix de l'un des mots exige une adresse sur 2 bits. S'il s'agissait de 8 mots de 2 bits organisés de la même façon, il faudrait un adressage sur 3 bits, le dernier pouvant tout simplement commander les amplificateurs de lecture ou d'écriture, etc.

De ce fait, le synoptique réel et pratique d'une RAM peut être tel que celui de la figure 2 où est symbolisée la MM 5262 de NS, qui est une RAM de 2 048 mots de 1 bit. L'entrée notée « Sélection de circuit intégré » est utilisée lorsque la mémoire globale regroupe plusieurs circuits intégrés du même type; ne seront alors *activés* (autorisés à fonctionner) que ceux qui reçoivent à cette broche le niveau logique approprié, 1 ou 0; les autres resteront inhibés. On procède ainsi à l'extension du réseau-mémoire à volonté.

3. LES CARACTÉRISTIQUES DES MÉMOIRES

Une fois défini le type de mémoire, RAM, ROM..., on considère :

- *Sa capacité* : c'est le volume d'informations, en bits, qu'elle peut emmagasiner. Les RAM de 4 K se sont généralisées en 1975, précédant les 16 K bits de 1977. En ROM, les 16 K étaient courantes dès 1975. La progression s'est ainsi adaptée à la pratique du binaire, avec une tendance à « sauter » les étapes des 2 K et 8 K...

- *Son organisation* : une mémoire travaille sur des mots dont il faut préciser la longueur et la quantité. Le produit nombre de mots par longueur du mot redonne sa capacité. Ainsi, on trouvera des ROM de 4 K bits fournissant soit 4 K mots de 1 bit, soit 1 K mot de 4 bits, soit 512 octets...

A ce propos, on peut noter que parmi les premières RAM apparues vers 1967 en bipolaire, figuraient des 16 bits réalisées sur 3 mm² de silicium. En 1974, les 4 K n'exigent que 16 mm² de silicium en MOS, ce qui témoigne des progrès réalisés par l'intégration. La densité s'accroît encore avec les technologies plus élaborées.

Les temps caractéristiques sont :

- *le temps d'accès* : c'est le temps écoulé entre le moment où l'adresse est appliquée et celui où l'information est présente à la sortie pour être lue.

- *le temps de cycle* : après une opération de lecture ou d'écriture, la mémoire peut avoir besoin d'un temps de réinscription (tores, par exemple) ou de « récupération ». Le temps de cycle est alors la somme de ce temps et du temps d'accès.

A partir de ces notions, on peut définir une *fréquence de travail maximale* (celle que la mémoire ne saurait dépasser), ou *typique* (conseillée en fonctionnement normal), le cadencement des opérations assurées par la mémoire pouvant être soumis à une horloge.

Dans le cas de registres dynamiques, MOS, CCD ou autres, on définira également une *fréquence minimale* de travail puisque ces éléments n'acceptent pas le « continu » (situation statique, ici), ou même des fréquences relativement basses.

La consommation des mémoires (courant ou puissance) est importante et s'exprimera éventuellement par :

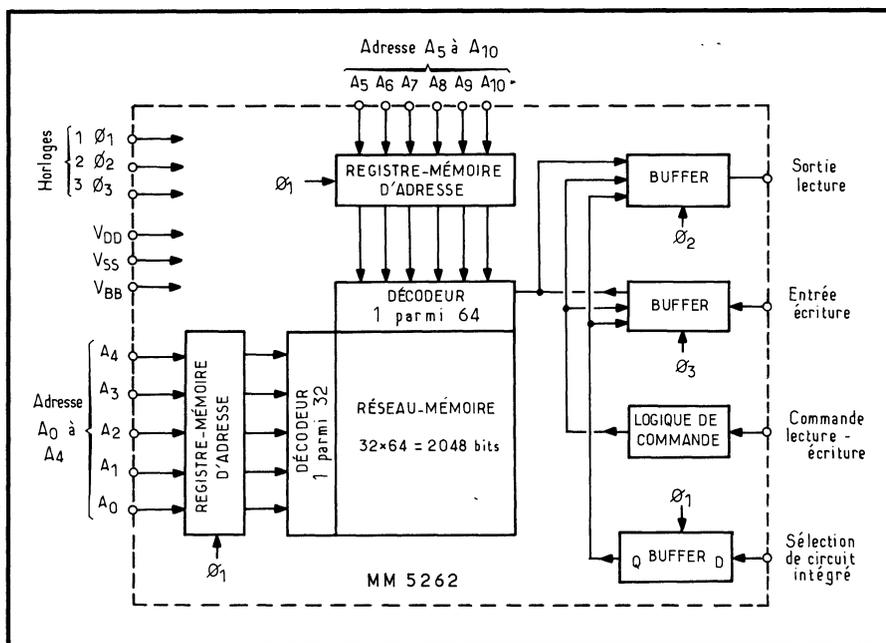


Fig. 2. — Synoptique d'une mémoire-type : la MM 5262 de National Semiconductor, RAM de 2048 mots de 1 bit.

- la consommation statique, ou au repos, la mémoire se trouvant hors des périodes d'exécution d'ordres;

- la consommation dynamique, ou en activité, en période d'exécution d'ordres.

Cette consommation pourra être évaluées par cellule ou pour l'ensemble de la mémoire, donc avec ses décodeurs amplificateurs, etc.

4. PRINCIPE DES ROM

La ROM la plus simple est constituée par un réseau de diodes, selon la disposition de la figure 3. Le réseau est, ici, de 4 mots de 4 bits (pour la commodité du dessin). Pour interroger la ROM, on ferme l'un des 4 interrupteurs, ce qui correspond à l'application d'une adresse via un décodeur à 1 parmi 4; la rangée concernée est mise sous tension et, chaque fois que l'on trouve une diode à une

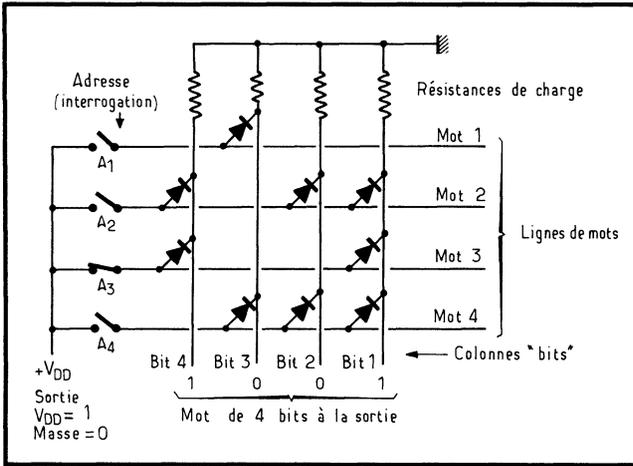


Fig. 3. — La matrice ROM à diodes : principe.

Fig. 4. — Matrice de 8 mots de 2 bits à diodes.

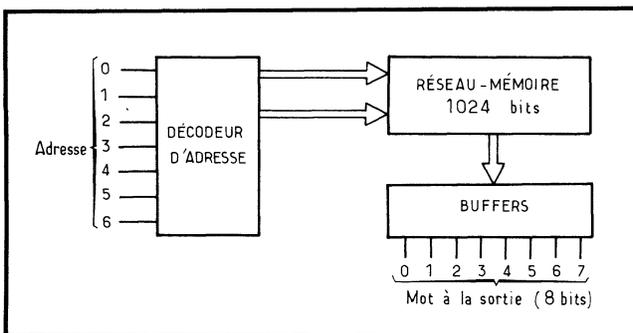
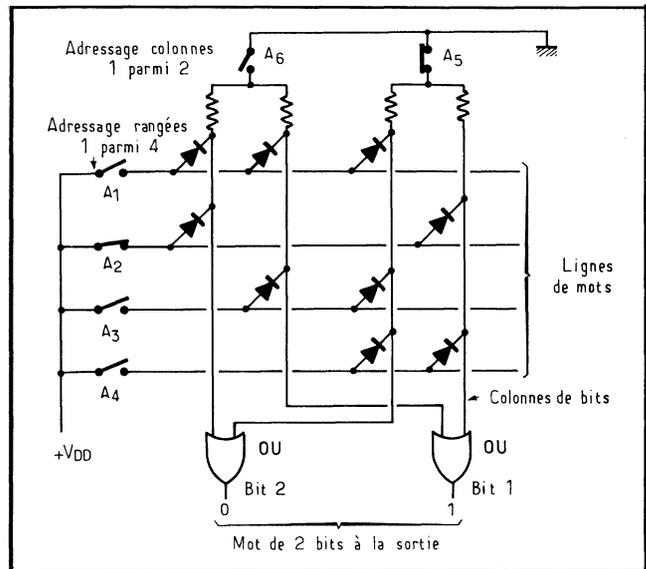


Fig. 5. — Synoptique type d'une ROM.

intersection avec une colonne, celle-ci fournira en sortie la tension V_{DD} , c'est-à-dire un 1 logique. Ici, le mot de sortie est donc 1001.

Les diodes peuvent être remplacées par tout autre élément unidirectionnel ou par un élément de couplage tel qu'un transistor. En effet, si l'on réunissait simplement et directement colonnes et rangées, on n'obtiendrait qu'un immense court-circuit.

Le même réseau peut se transposer en matrice de 8 mots de 2 bits, à la condition de consentir en plus à un adressage « colonnes » (fig. 4).

Le plus souvent, on ajoute sur les lignes de sorties des amplificateur-buffers, et de ce fait, le synoptique de principe d'une ROM revient celui de la figure 5 où l'adressage a été regroupé en un mot de 7 bits, pour lire ici une mémoire de 1 024 bits au format de 128 mots de 8 bits.

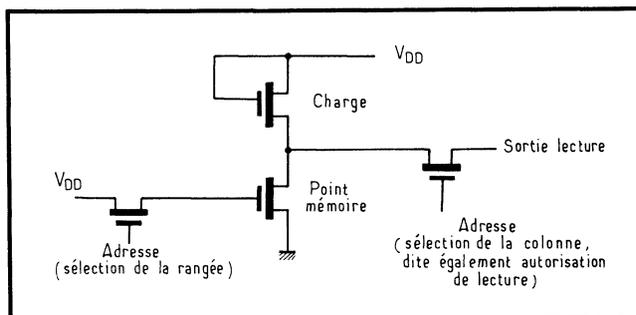
A la place des diodes peuvent figurer des MOS à chaque point de croisement. La tension d'interrogation est alors appliquée sur les gates et l'on peut songer, au moins, à deux formules pour stocker des 0 et des 1 :

1. Comme avec les diodes, le MOS est physiquement présent ou absent.

2. Ou alors, tous les MOS existent mais certains possèdent un gate normal, avec oxyde mince, alors que d'autres ont le métal de gate déposé sur une couche d'oxyde épais, ce qui équivaut pratiquement à rendre impossible le déblocage du MOS.

Dans les deux cas, le résultat est rigoureusement semblable, d'ailleurs. La résistance de charge commune est, ici un MOS, l'adressage pouvant être assuré également via des MOS (fig. 6). La structure-type d'une ROM à MOS, de 32 bits

Fig. 6. — Principe de l'adressage du MOS-mémoire.



en réseau de 4 rangées par 8 colonnes, fournissant 16 mots de 2 bits, est représentée figure 7.

Pratiquement, les temps d'accès des ROM (qui sont leurs temps de lecture) varient entre 50 ns et 3,5 μ s selon type et complexité, et les réseaux atteignent jusqu'à 64 K bits (1978) pour une dissipation de l'ordre de 0,01 à 0,2 mW/bit.

5. LES PROM

Les PROM (*ROM programmable*) constituent une variante des ROM; mais alors que ces dernières sont enregistrées par le fabricant de circuits intégrés par

masquage (et donc pour de grandes séries), les PROM sont livrées non-inscrites à l'utilisateur final. Plus exactement, toutes les cellules se trouvent dans le même état, 0 ou 1, et c'est l'utilisateur qui inscrira l'état inverse là où il le désire, selon son programme. Cette inscription est irréversible et n'est plus effaçable; pour l'assurer, on dispose de plusieurs techniques :

1. *Par fusion de fusibles*, montés en série avec chaque point-mémoire;
2. *Par court-circuit d'une jonction* d'un transistor bipolaire, la seconde apparaissant alors comme une diode en service.

Ces deux technologies ne s'appliquaient qu'aux bipolaires jusqu'en 1978; faisaient exception quelques réalisations en MOS complémentaires, à fusibles (produits de *Harris*). Pour mémoire, on notera encore une troisième possibilité :

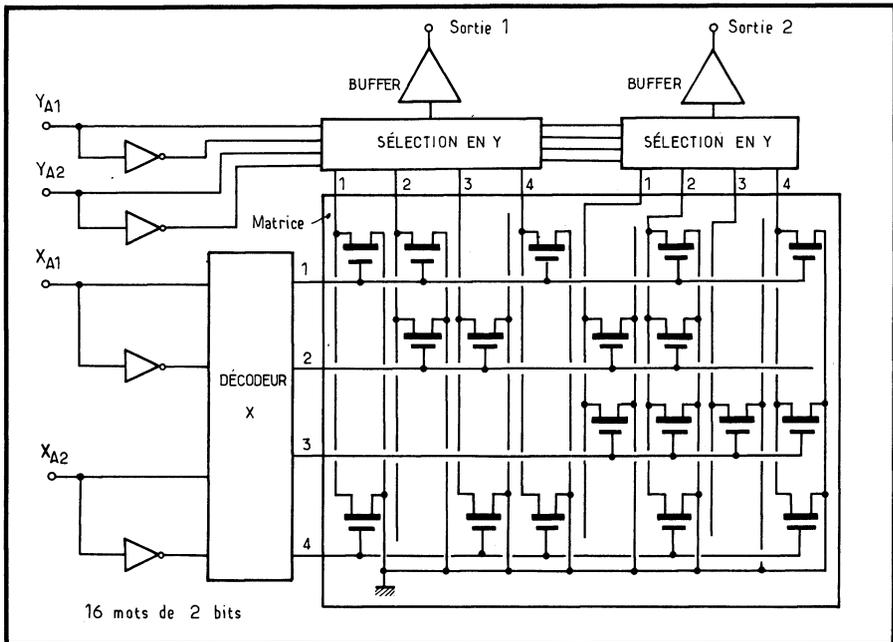


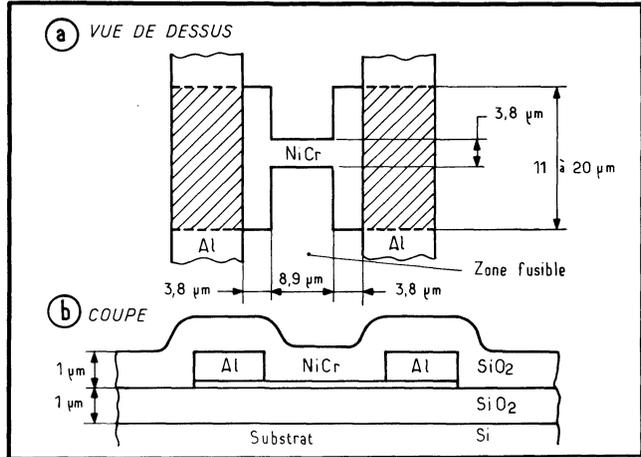
Fig. 7. — Schéma d'une mémoire de 16 mots de 2 bits.

3. *On utilise des EPROM* (dont on verra ce qu'elles sont plus loin) mais on leur ôte toute possibilité d'effacement en supprimant la fenêtre transparente aux ultraviolets qui servent à décharger les transistors-mémoires. Le boîtier est alors bien plus économique, et peut être en plastique.

PROM à fusible

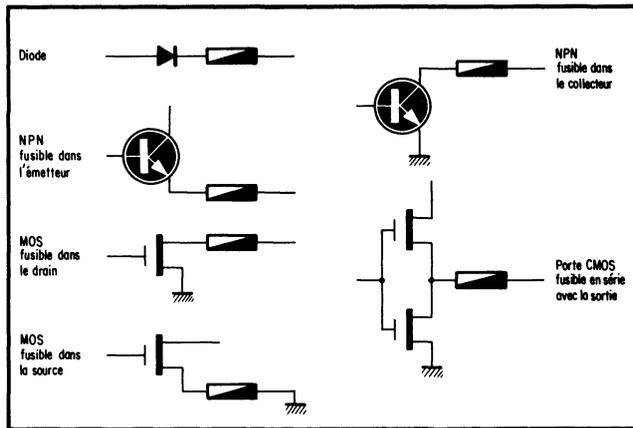
Dans les PROM à fusible, on insère un véritable fusible intégré en série avec la diode classique des ROM ou le transistor qui établit la liaison colonne-rangée.

Fig. 8. — Structure du fusible d'une PROM.



Il s'agit simplement d'une connexion dont on réduit fortement la largeur sur une petite portion de son parcours (fig. 8); elle est en aluminium, nichrome ou même silicium polycristallin. Une forte impulsion de courant suffit pour volatiliser et

Fig. 9. — Quelques structures de points-mémoires pour PROM.



détruire cette zone étroite, et donc pour isoler le composant en série. Celui-ci est donc classiquement une diode ou un transistor bipolaire, mais pourrait être un MOS, selon les quelques configurations données figure 9.

Dans le cas d'un réseau à diodes (fig. 10 : dans ce réseau *Harris*, seules quelques diodes avec leur fusible ont été dessinées), la programmation s'effectue en adressant une rampe de courant au point-mémoire visé : jusqu'à 750 mA. L'élévation de température provoquée au niveau du fusible le fait fondre; ici, il est en aluminium et le métal se rétracte alors de part et d'autre d'une coupure centrale franche.

Dans le schéma, on voit quel dispositif « manuel » est utilisé à cette fin : avec la position indiquée de l'interrupteur S_A , le condensateur C_1 est déchargé à la masse; les interrupteurs passant en position (2), le condensateur se charge et pilote le Darlington qui délivre une rampe de courant à la mémoire, jusqu'à destruction du fusible. Un temporisateur (non dessiné) peut surveiller le temps nécessaire à l'opération. Ramené en position (1), l'interrupteur S_B sert à vérifier que le circuit a bien été coupé car le voyant ne doit alors pas s'allumer.

Si c'est de l'aluminium qui constitue le fusible, sa fusion survient à $650\text{ }^\circ\text{C}$ environ et n'affecte pas donc le circuit intégré, protégé par une couche de dioxyde, dont la température de fusion est $1\ 350\text{ }^\circ\text{C}$.

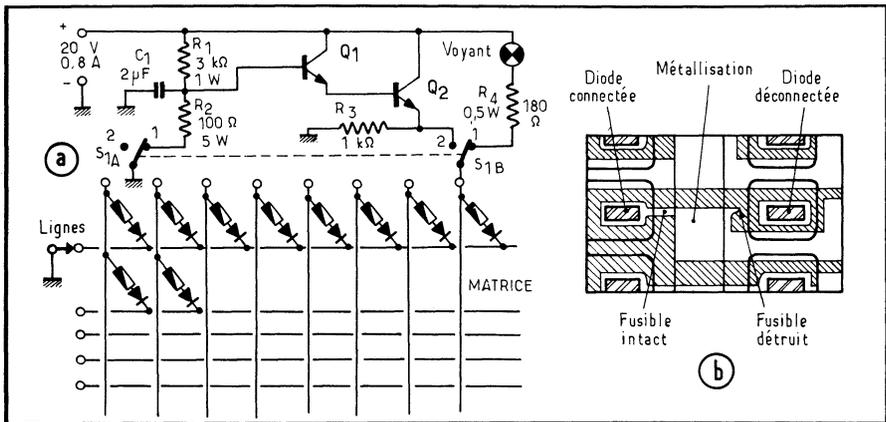


Fig. 10. — Un réseau de PROM à diodes et sa programmation (a). Deux diodes : l'une est restée connectée, l'autre est isolée car son fusible a été détruit (b); les métallisations apparaissent en clair.

Dans le cas d'un transistor bipolaire, l'impulsion de courant est moindre : 15 à 30 mA si le fusible se trouve dans le circuit émetteur, l'adressage se faisant sur la base, et ce pendant quelques 200 ms (exemple : la 0512 de Harris).

Cette même société a proposé des PROM en MOS complémentaires, dont l'intérêt essentiel est la très faible consommation : $500\ \mu\text{W}$ au repos pour une PROM de 1 024 bits et 50 mW en activité. Il s'agit de mémoires avec fusibles au nickel-chrome.

PROM à court-circuit de jonction

Une autre méthode, typiquement bipolaire, a été proposée par *Intersil*; elle consiste à réaliser deux jonctions tête-bêche, qui bloquent par conséquent la ligne sur laquelle elles sont montées en série. Ces deux jonctions sont celles d'un transistor dont la base n'est pas connectée. Une forte impulsion fait partir en avalanche une jonction et le métal du contact, de l'aluminium, va migrer le long de la ligne de courant jusqu'à court-circuiter cette jonction; l'avalanche cesse alors d'elle-même et il ne reste plus qu'une diode (la seconde jonction) en service. On est ramené au cas d'une ROM à diode.

Ce processus a reçu le nom de technologie à *migration par avalanche induite*, soit AIM (de « *avalanche induced migration* »).

6. EPROM ET REEPROM

Cette troisième catégorie de mémoires mortes comprend des éléments dont l'appellation synonyme est :

- EPROM, pour *mémoire PROM programmable électriquement* ou :
- REEPROM, pour *PROM reprogrammable par l'utilisateur*.

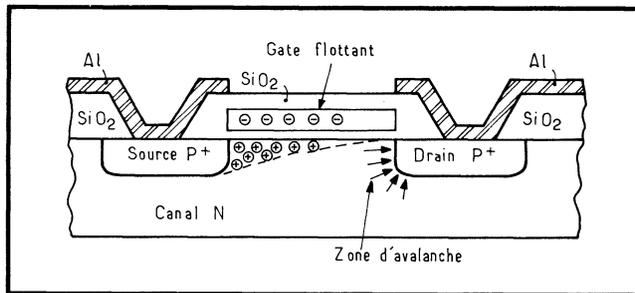
Comme les précédentes (les PROM), les EPROM sont livrées non inscrites à l'utilisateur qui les enregistrera individuellement; mais il pourra encore, si besoin est, effacer le programme inscrit et recommencer un nouvel enregistrement. Parce que l'effacement et l'inscription sont des processus relativement longs (comparés à la lecture), on dit encore de ces mémoires qu'elles sont à *lecture majoritaire*, ce qui se note RMM (*read mostly memory*). Les principales techniques qui prévalent sont :

1. Les EPROM à inscription « électrique » et effacement par ultraviolet, les plus usuelles jusqu'en 1978.
2. Les EAROM ou EEROM, recourant aux transistors MNOS, dont le développement est plus récent et qui portent les plus grandes promesses.
3. S'y ajoutent les mémoires à verre semiconducteur, qui ont eu leur temps de succès relatif et semblent actuellement abandonnées.

La EPROM à effacement par ultraviolet

La EPROM à effacement par ultraviolet utilise un transistor MOS spécial, dit à *stockage de charge*, recourant à un gate flottant. En effet, son gate est noyé dans l'oxyde sans aucun contact extérieur (fig. 11), la silice sous-jacente mesurant 100 nm d'épaisseur contre 1 000 nm au-dessus.

Fig. 11. — Principe du MOS à gate flottant.



Si l'on applique une impulsion fortement négative au drain de ce PMOS, la jonction PN constituée par le drain et le substrat part en avalanche et, par effet tunnel, des électrons sont injectés dans le gate flottant. Leur nombre dépend de l'amplitude et de la durée de l'impulsion. Lorsque celle-ci disparaît, les électrons restent piégés et provoquent la création d'un canal P dans le substrat, donc rendent le MOS conducteur.

Selon la firme *Intel*, cette charge se conserverait plus d'un siècle. Pour l'effacer on utilise un rayonnement ionisant, de l'ultraviolet qui permet au gate de se décharger. Il faut donc que le MOS soit doté d'une fenêtre transparente aux UV, en quartz.

Ce transistor s'appelle le FAMOS, de *floating avalanche injection MOS*. Il a mené des mémoires complètes dont l'une des premières est la 1701, de 2 048 bits en réseau de 256 mots de 8 bits (fig. 12). La lecture s'effectue en 700 ns, à une fré-

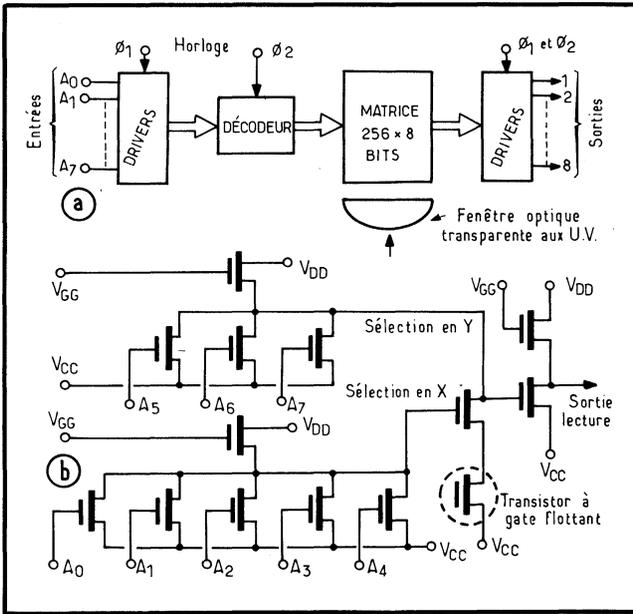


Fig. 12. — La REPRM 1701 de *Intel*; synoptique (a) et schéma électrique partiel (b).

quence maximale de 1 MHz pour une consommation moyenne de 30 mA (alimentation de ± 5 V et -9 V). L'effacement est assuré avec une source d'UV à $2\,537 \text{ \AA}$, avec une intensité de 10 mW/cm^2 , située à 4 cm du CI. Dix minutes d'exposition sont plus que suffisantes.

EAROM et EEROM

Les EAROM, ou EEROM, procèdent de la même idée : la mémoire morte peut être enregistrée par l'utilisateur et éventuellement effacée pour recevoir un nouveau programme, mais ici l'effacement est électrique, tout comme l'inscription. Ces sigles signifient :

- EAROM : *electrically alterable ROM*, soit ROM altérable électriquement;
- EEROM : *electrically erasable ROM*, soit ROM effaçable électriquement.

Ils sont souvent utilisés l'un pour l'autre bien qu'en toute rigueur, ils ne sont pas synonymes.

Les EAROM sont très intéressantes en raison de leur traitement totalement électrique. Elles pourraient supplanter les EPROM dans toutes leurs applications et en outre, servir dans des systèmes tels que caisses enregistreuses pour stocker les totaux d'un jour sur l'autre, ou en informatique mémoriser des résultats intermédiaires, des adresses de départ de tables, de boucles ou de programmes utilitaires; jouer le rôle de relais à l'état solide avec mémoire; stocker des caractéristiques industrielles en commande numérique, par exemple, etc. De plus, et dès que leur vitesse aura été améliorée (réduction des temps) et rejoindra celle des RAM classiques, elles pourraient prendre la place de celles-ci en se comportant en RAM non volatiles.

Ces EAROM sont fondées sur un transistor appelé MNOS dont la structure, résumée par son sigle (métal-nitride-oxyde-silicium) est donnée figure 13. Une couche très mince de silice, de moins de 2 nm, est déposée sur le silicium du substrat et prise en sandwich entre ce substrat et du nitride de silicium.

Fig. 13. — Le transistor MNOS.

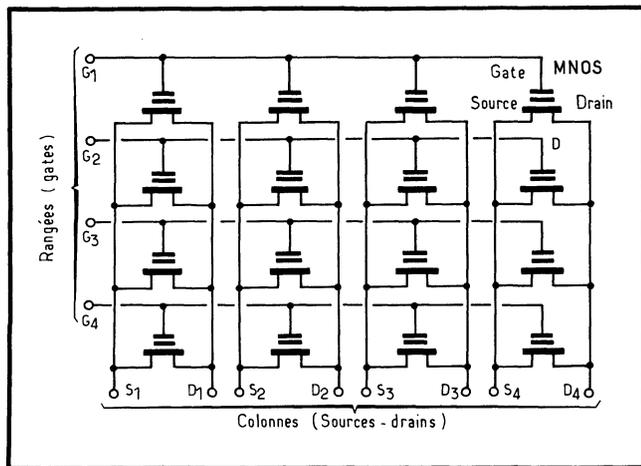
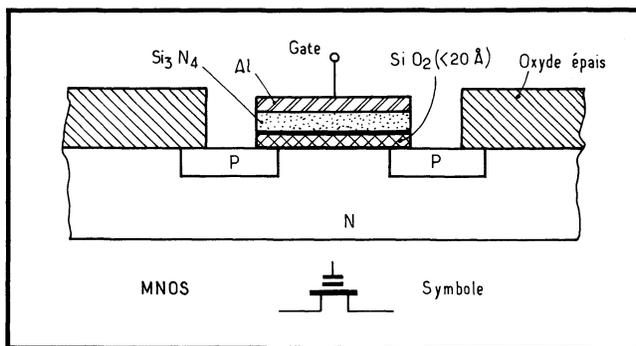


Fig. 14. — Réseau-mémoire à MNOS.

Son fonctionnement est le suivant : si l'on polarise positivement le gate, avec des tensions relativement importantes (les drains et source étant normalement alimentés), des électrons traversent l'oxyde par effet tunnel et sont piégés à l'interface oxyde-nitride; ils y restent lorsque la polarisation disparaît. Cette interface joue désormais le rôle d'un gate annexe, et maintient le MOS conducteur.

A l'inverse, et si l'on applique une forte tension négative, les électrons piégés sont chassés et, toujours par effet tunnel, traversent en sens inverse la silice. Le MOS reste alors bloqué. On dispose donc d'un MOS *programmable et effaçable électriquement*, avec deux tensions de seuil, V_{T+} et V_{T-} .

Les premiers MNOS ont été commercialisés sous forme de transistors discrets; puis sont venus les circuits intégrés, en réseau de 4 par 4 tout d'abord (fig. 14), puis en configurations plus complexes de plusieurs kilobits avec circuits de commandes inclus, tels ceux de *General Instruments*, ou de *EFCIS* pour la France, par exemple.

Structures de EFCIS en SOS

Le transistor MNOS de *EFCIS* (*Société pour l'Étude et la Fabrication de Circuits Intégrés Spéciaux*, implantée à Grenoble) a la structure de la figure 15. Il est réalisé sur substrat isolant, corindon (ou saphir). L'oxyde fin a une épaisseur de 3 nm et le nitrure de 50 nm.

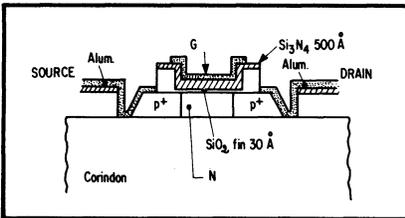


Fig. 15. — Le transistor MNOS de EFCIS.

Lors de l'application d'une tension positive sur le gate et à partir d'une tension de seuil d'injection (le substrat étant connecté à la masse), des électrons franchissent par effet tunnel la barrière constituée par la fine couche d'oxyde de silicium. Ces électrons sont alors stockés dans les pièges situés à l'interface oxyde-nitrure de silicium. Par application d'une tension négative sur le gate, on peut ensuite vider les pièges et retrouver la situation électrique initiale. La présence ou l'absence des électrons à l'interface détermine les tensions de seuil entraînant la conduction du transistor; on obtient ainsi

- Pour un MOS à *enrichissement* (absence d'électrons), une tension de seuil V_{TH} de l'ordre de -3 V.

- Pour un MOS à *déplétion* (présence d'électrons), une tension de seuil V_{TH} de l'ordre de $+3$ V.

Cette tension de seuil garde cette même valeur après suppression de la tension de programmation, ce qui est indispensable pour un fonctionnement non volatil et évolue ainsi comme le montre la figure 16 pour les deux types de transistors.

EFCIS a ainsi commercialisé dès 1977 une mémoire de 1 024 bits organisée en 128 mots de 8 bits, entièrement décodée et compatible TTL. Les MNOS sont effaçables par bloc de 1 kilobit en 100 ms, et programmables par mots de 8 bits en 100 ms également. Le temps de cycle, égal au temps d'accès, est inférieur à 1 μ s et la dissipation est inférieure à 400 mW (écriture) ou 150 mW (lecture), pour des tensions d'alimentation de 5 V et -12 V en lecture, ou $+5$ et -30 V en écriture. La durée de rétention de l'information dépasse 5 ans et le nombre de cycles écriture-effacement excède 10^5 . Le circuit est présenté en boîtier DIL à 24 broches.

Fig. 16. — Évolution de la tension de seuil du MNOS de EFCIS en fonction de la tension gate-substrat.

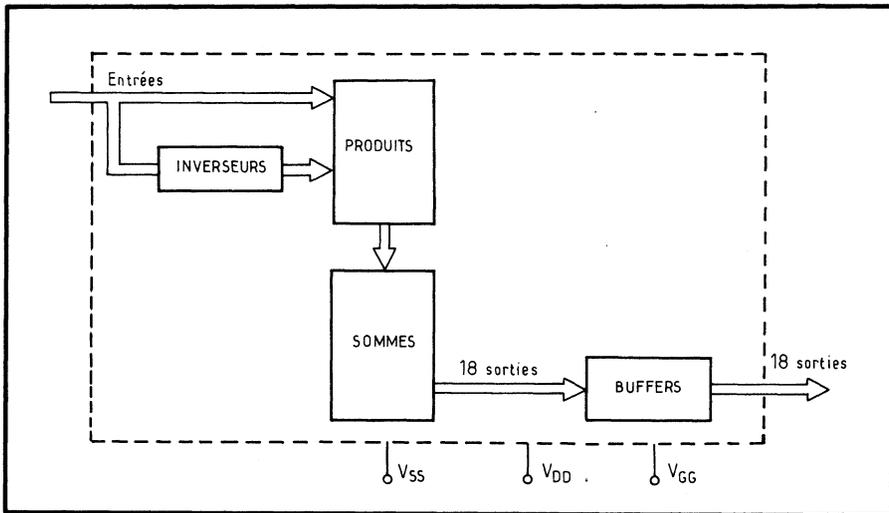
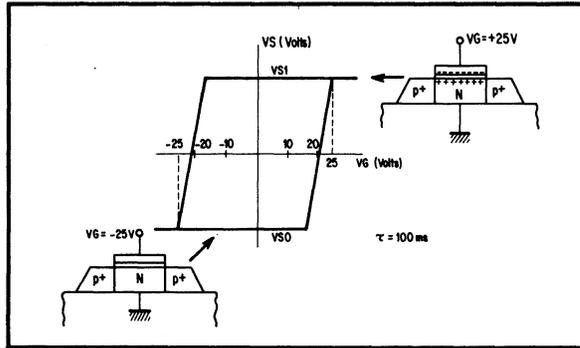


Fig. 17. — Synoptique d'un réseau logique programmable.

7. LES RÉSEAUX LOGIQUES PROGRAMMABLES

Le réseau logique programmable, ou PLA (de *programmable logic array*) constitue une façon élégante et souvent économique de réaliser des fonctions logiques complexes. En effet, des éléments communs peuvent alors servir plusieurs fonctions. D'autre part, le réseau logique est produit par le fabricant jusqu'au dernier masquage, qui ne sera exécuté que sur les instructions de l'utilisateur (tout comme l'on programme les ROM à la demande). De plus, on peut réaliser des PLA programmables par l'utilisateur, comme on le fait pour les PROM.

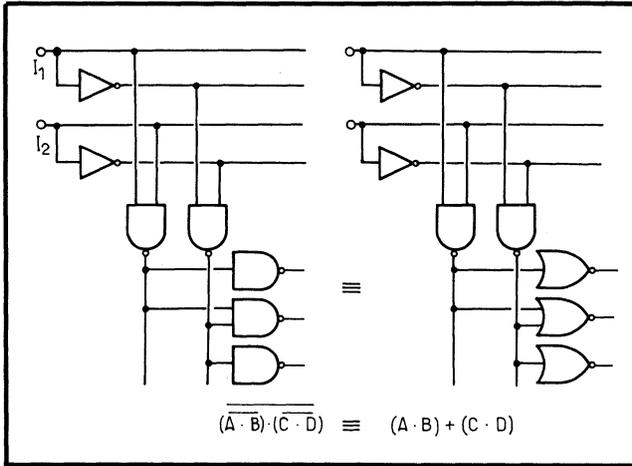


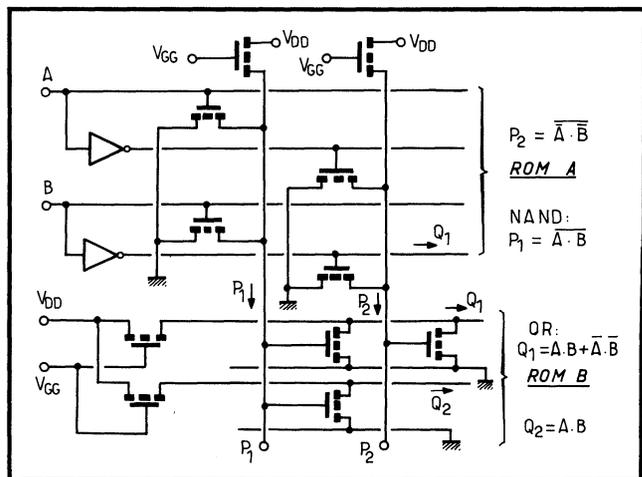
Fig. 18. — Équivalent logique.

La structure d'un PLA est simple; en logique, une fonction peut se ramener à une somme de produits ou à un produit de sommes. Si l'on choisit la somme des produits :

- une première matrice fournit les produits (matrice AND);
- une seconde génère leur somme (matrice OR).

Le synoptique d'un PLA type est donné figure 17. Les matrices AND et OR sont constituées par des portes logiques. L'équivalent logique général du PLA est dessiné figure 18. Réduit aux MOS réellement présents, le schéma de la figure 19 montre un autre type de réseau.

Fig. 19. — Exemple de réseau logique à MOS.



La logique combinatoire est aisée à réaliser avec une logique classique, ou des ROM. Dans ce cas cependant, chaque nouvelle variable d'entrée double la capacité de mémoire nécessaire. Ce n'est plus le cas des PLA.

Ainsi, le PLA type TMS 2000 réalise 60 fonctions sur 25 variables; en ROM, cela exigerait une capacité de 225 mots (17 entrées « externes » et 8 entrées de flip-flop), soit 8 288 608 mots de 26 bits (18 sorties « externes » et 8 pour les flip-flops), soit au total 218 103 808 bits! On reconnaît là tout l'avantage des PLA.

Dans le cas de la logique séquentielle, les séquences sont assurées en réinjectant à l'entrée des signaux de sortie selon la combinaison souhaitée. Avec le PLA cette action est assurée à l'intérieur même du circuit, ce qui économise d'autant en bornes d'accès.

CHAPITRE XV

LES MÉMOIRES VIVES

Les mémoires vives, ce sont les RAM, qui peuvent être lues et ré-inscrites à volonté sans pénalisation en temps. Ce sont également des mémoires adressables, ce qu'indique leur sigle : random access memory, soit mémoire à accès aléatoire (dont on peut adresser directement une cellule quelconque). Elles existent en bipolaire et en MOS, et se manifestent depuis peu en CCD et en bulles magnétiques (MBM : mémoire à bulles magnétiques).

1. LA PRODIGIEUSE ÉVOLUTION DES CELLULES RAM

Encore toute récente, l'évolution de la structure des cellules-mémoires MOS a quelque chose d'édifiant. Leurs principes ont été décrits antérieurement, mais on les développera mieux ici, qu'il s'agisse de :

- *flip-flops*, menant à des mémoires dites *statiques* parce qu'elles peuvent conserver indéfiniment leur état sans intervention extérieure (tant que l'alimentation subsiste, bien entendu);
- *capacité intrinsèque* du MOS, utilisée dans des cellules *dynamiques* ainsi appelées parce qu'elles n'acceptent pas un fonctionnement à la fréquence nulle, à moins d'être périodiquement rafraîchies.

Brossée succinctement, l'évolution des cellules RAM a ainsi été la suivante :

1. La cellule de base d'une mémoire a été, dès l'origine (vers 1967!), un flip-flop d'abord bipolaire, donc comportant 2 transistors, 2 résistances et au moins 2 transistors d'adressage ou de lecture écriture (fig. 1 *a*), soit au moins 6 composants.

2. Transposée en MOS monocanaux, elle mène à la structure à 6 MOS (en *b*), très satisfaisante.

3. Les lignes d'écriture et de lecture pouvant être multiplexées, on supprime alors un MOS pour passer à la cellule à 5 transistors (*c*).

4. Pour simplifier encore la cellule, il faut passer à la formule semi-dynamique MOS de la figure *d*. Il ne subsiste du flip-flop que Q_1 et Q_2 , et ce sont les capa-

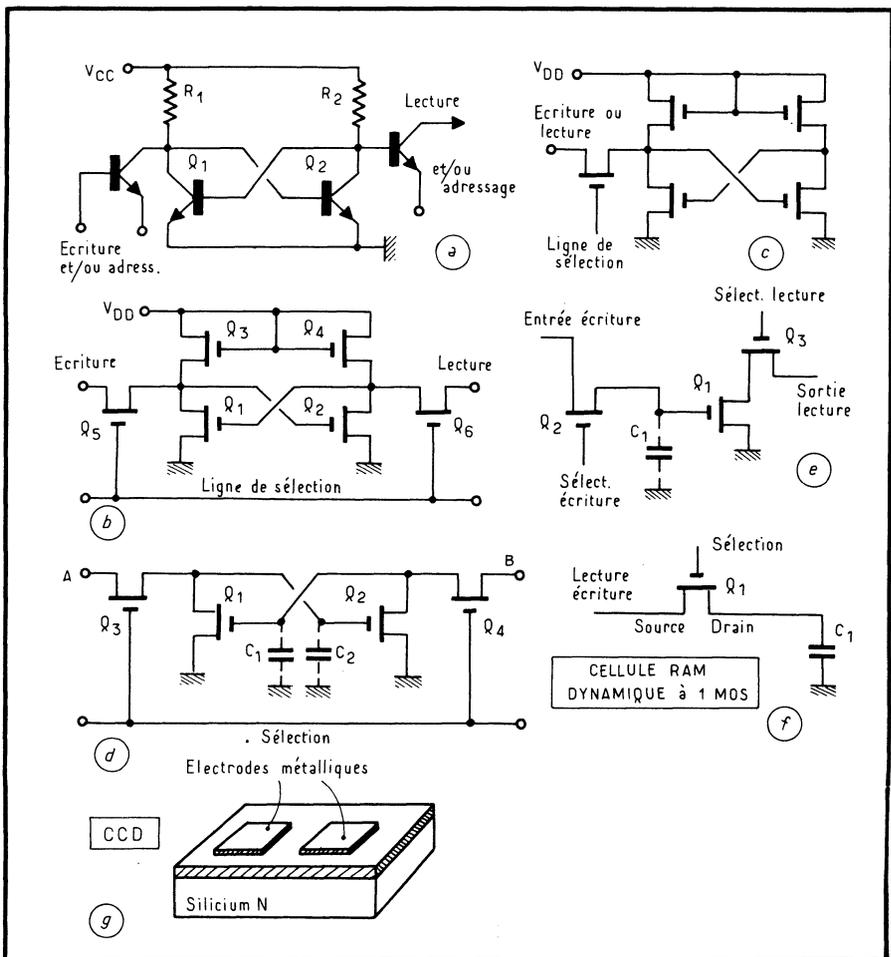


Fig. 1. — Quelques cellules de RAM.

cités de structure C_1 et C_2 qui mémorisent l'information. Les MOS Q_3 et Q_4 servent à la sélection de la cellule RAM voulue, en position *écriture* ou *lecture*.

5. La structure de flip-flop n'étant plus indispensable en dynamique, on peut encore supprimer un transistor et passer à 3 (en e). La mémorisation s'effectue grâce à la capacité de structure C_1 de Q_1 , les MOS servant à la sélection et à la lecture-écriture.

6. On peut encore se servir de la capacité drain-substrat du MOS et, dans ce cas, ramener la cellule à un seul transistor (en).

7. Puisqu'en réalité, seule la capacité de structure intervient, on peut totalement se passer de MOS pour le réseau-mémoire, et concevoir une cellule ne comportant qu'un condensateur. Il ne reste plus alors qu'à passer aux CCD (g).

8. Enfin, avec les bulles magnétiques dont l'étude n'entre pas dans le cadre de cet ouvrage, ce sont des cylindres d'aimantation inverse qui servent à mémoriser l'information.

2. LES RAM STATIQUES A MOS MONOCANAUX

La cellule complète à 8 transistors

La cellule statique de base, à flip-flop, regroupe 8 transistors si on lui adjoint les MOS pour l'adressage (fig. 2). Son fonctionnement est le suivant.

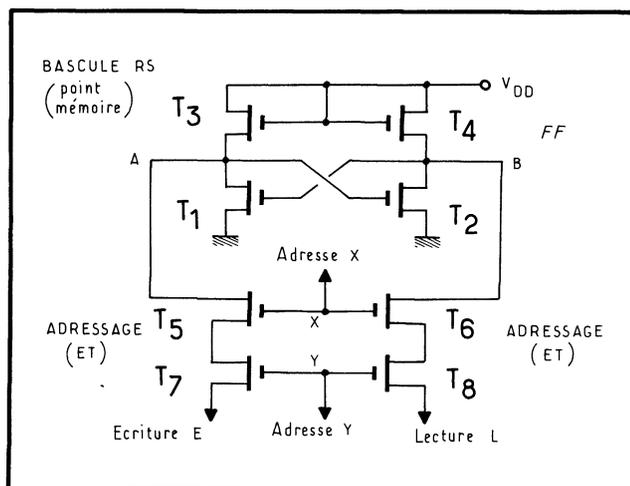


Fig. 2. — Point-mémoire complet à 8 MOS.

Le décodeur X active la ligne X et rend conducteurs T_5 et T_6 ; le décodeur Y en fait autant avec T_7 et T_8 ; les entrées A et B, de lecture et/ou d'écriture sont alors connectées à la bascule.

On constate cependant que les portes AND (ET) d'adressage viennent en parallèle sur les charges, comme le montre la même figure redessinée (fig. 3), et

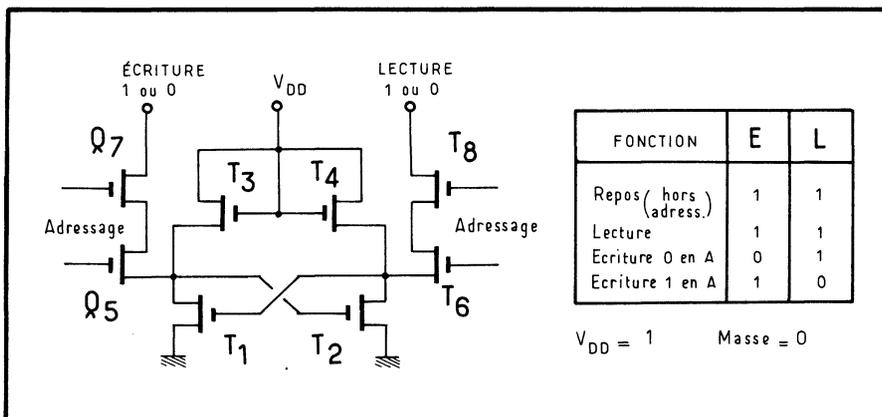


Fig. 3. — Fonctionnement de la cellule MOS statique.

constituent avec elles un diviseur de tension qui fait que le courant est différent à la lecture et à l'écriture, pour une même ligne d'accès.

Une telle cellule occupe quelque 0,03 mm² de silicium, ce qui est encore moins que les 0,04 mm²/cellule de la TTL type 7489. La dissipation est de l'ordre de 1 mW/bit, contre environ 6 mW/bit en TTL.

Cellule classique à 6 transistors : la 1101

En fait, les doubles transistors d'adressage créent une redondance superflue, ce qui permet de ramener la cellule à 6 MOS (fig. 4). Q₁ et Q₂ constituent le FF chargé par Q₃ et Q₄; si Q₅ et Q₆ sont bloqués, le flip-flop est isolé et conserve en mémoire la dernière position acquise. Supposons que Q₁, Q₃ fournissent l'état haut (1) et Q₂, Q₄ l'état 0.

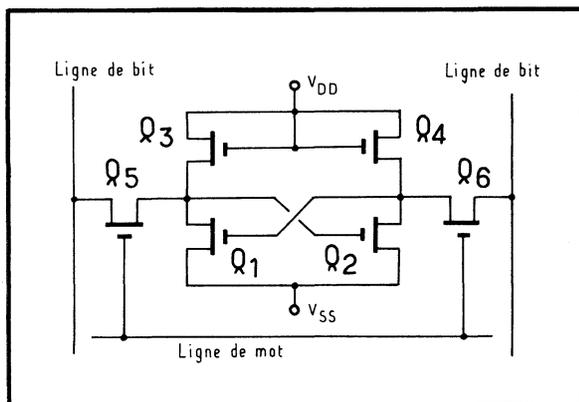


Fig. 4. — Cellule statique à 6 MOS.

Pour lire cette information, Q_5 et Q_6 sont débloqués grâce à une impulsion appliquée sur la ligne de mot. Un courant en témoignera via Q_6 dans la ligne de bit de droite, qu'on peut appeler ligne de bit 1, par opposition à celle de gauche, qui sera la ligne de bit 0. L'écriture d'une information se fait également en rendant passants Q_5 et Q_6 et en portant l'une des lignes de bit à l'état bas (0) : elle impose son état à l'inverseur considéré, donc au flip-flop, après toutefois que la capacité de structure chargée du MOS se soit déchargée.

Une cellule de ce type a été adoptée, par exemple, par Intel pour les classiques 1101/11011; cette mémoire 256 bits consomme 2 mW/bit en activité, et 50 μ W/bit au repos; son accès est assuré en moins de 800 ns.

Cellule 6 MOS à charge à déplétion (type 2102 A)

En recourant à des charges à déplétion (et non plus classiques à enrichissement), Intel a réalisé la 2102 A, en NMOS, dont l'avantage principal sur la 2102

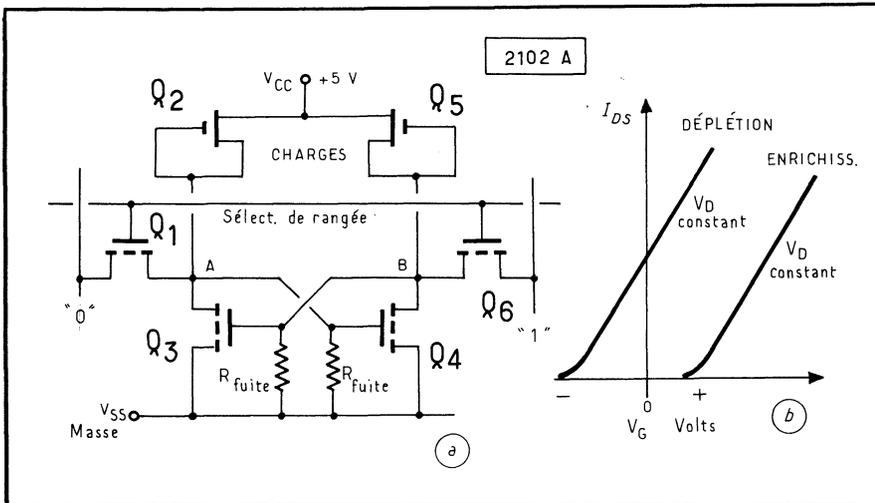


Fig. 5. — Cellule statique avec charges à déplétion (a), et courbes caractéristiques (b).

normale (à charges à enrichissement) est de permettre une réduction de puissance consommée : 42 mW maximum à l'état « faible consommation » contre 150 mW typiques normalement.

La cellule mémoire a exactement la même structure (fig. 5). Supposons que Q_6 applique un 1 au point B, donc au gate de Q_3 ; ce MOS conduit et débite dans Q_2 , charge à déplétion mais d'impédance plus élevée que Q_3 ; la tension en (a) diminue alors, et ce d'autant plus qu'au fur et à mesure que la tension baisse, la commande de la charge Q_2 diminue, ce qui accroît d'autant son impédance; A passe ainsi à V_{SS} , ce qui bloque Q_4 ; la charge Q_5 restant conductrice, B évolue vers le V_{CC} (au 1), ce qui accroît jusqu'au niveau maximal la commande de Q_5 et réduit son impédance : B passe à V_{CC} . Le seul courant fourni par Q_5 est le courant de fuite R_L .

Puisque les charges à déplétion sont normalement conductrices, qu'elles ne provoquent pas de chute de tension (égale à la tension de seuil avec les types à

enrichissement) et qu'une tension négative (pour le NMOS) seule peut les bloquer, on peut réduire l'alimentation jusqu'à une tension dont la limite est celle du bon fonctionnement de la bascule Q_3 et Q_4 , soit 1,5 V (au lieu de 5 V), ce qui permet effectivement de réduire la puissance consommée.

Bien entendu, un montage de commande est ici à prévoir et il comporte certaines servitudes aisées à surmonter. Cette fonction à basse consommation sera utile, en particulier, pour relayer une défaillance de l'alimentation fournie par le secteur (fig. 6); si, au temps 0, l'alimentation principale fait défaut, le condensateur C en se déchargeant prolonge le cycle de travail; lorsque la tension V_L aux

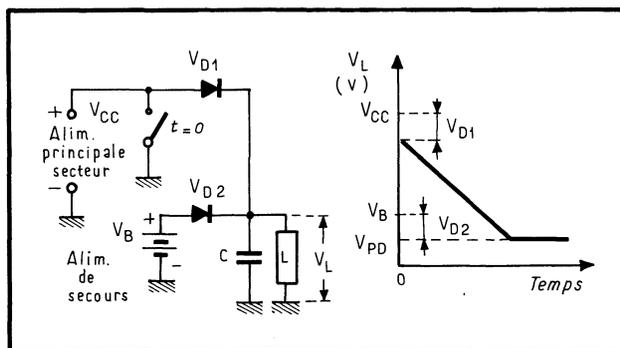


Fig. 6. — Le fonctionnement en « faible consommation » permet l'alimentation de secours par batterie.

bornes de la charge est, à une chute de tension de diode près, égale à la tension V_B de la batterie, cette dernière prend le relais. Lors du rétablissement de l'alimentation principale, la tension croissant aux bornes du condensateur bloquera la diode D_2 et isolera à nouveau la batterie.

La gamme importante des mémoires statiques se complète sans cesse, avec des types plus rapides ou plus denses. En 1978, on annonçait des temps d'accès de 35 ns et des capacités de 32 K bits.

3. RAM STATIQUES EN CMOS

En MOS complémentaires, la cellule mémoire statique est un flip-flop, composé de deux inverseurs MOS rétro-couplés. Avec ses deux transistors d'adressage, des MOS à canal P, elle présente l'aspect des figures 7 a et b.

Le circuit CD 4061 A de RCA (série 4000), par exemple, regroupe 256 cellules de ce type en réseau de 256 mots de 1 bit répartis, dans la matrice, en 16 rangées et 16 colonnes (en c). Lors de la lecture, l'état de la cellule est transféré à un second PMOS, puis à l'amplificateur de sortie (fig. 8 a). Ces deux PMOS en série permettent l'adressage en X et en Y.

Si, à la lecture, une seule voie est utilisée, à l'écriture on recourt aux deux : information directe et complémentée. L'adressage s'effectue à partir de 4 bits appliqués à un circuit de sélection d'une ligne parmi 16, répétés pour les deux axes de coordonnées; en plus, le décodeur d'adresse X est doté d'une borne d'*autorisation* qui inhibe le circuit entre deux opérations successives.

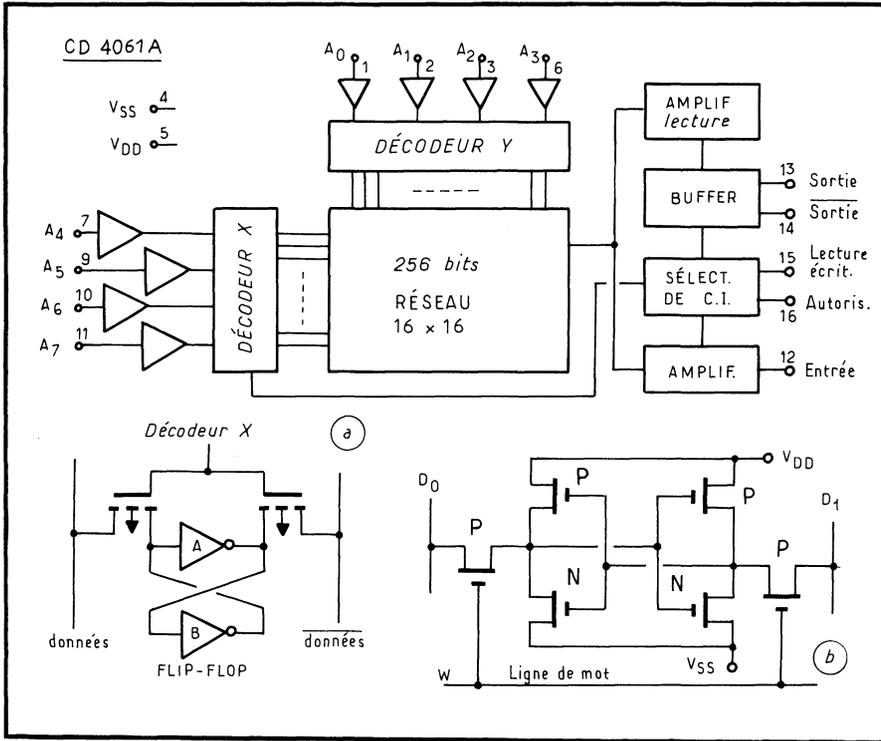


Fig. 7. — Cellule CMOS de base (a et b); organisation de la RAM type CD 4061 A de RCA (b).

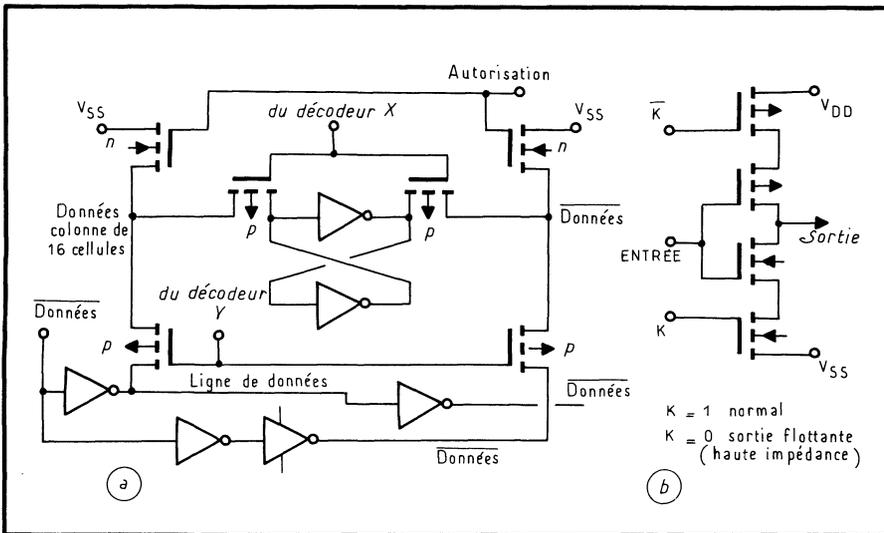


Fig. 8. — Détail des circuits d'accès de la CD 4061 (a) et structure de la porte 3 états (b).

On dispose également d'une entrée globale d'autorisation, qui sert à sélectionner la RAM utilisée lorsque plusieurs d'entre elles sont regroupées, et à décharger la ligne d'adressage Y entre deux opérations (puisque le décodeur Y est toujours en service).

Dernière caractéristique, enfin, des inverseurs trois états apparaissent pour la commande des bits ou sorties. Ces circuits 3 états sont le pendant de ceux qui s'appliquent à la TTL, mais ils sont bien plus simples ici (fig. 8 b) puisqu'il suffit de deux transistors en série avec la porte de sortie pour, si K est *haut*, la bloquer totalement et isoler la sortie du circuit intégré. Ainsi, les sorties peuvent-elles être assemblées en OU câblé.

Mémoire statique, cette RAM est dispensée de tout rafraîchissement. Ses entrées sont protégées par des diodes. Le circuit consomme $0,2 \mu\text{W/bit}$ au repos et $120 \mu\text{W/bit}$ avec cycles de $1 \mu\text{s}$. Il est alimenté par une unique tension, entre 3 et 15 V et sa sortie peut piloter 1 charge TTL. Le temps d'accès est de 290 ns. On notera en particulier que la consommation globale est des plus réduites, ce qui constitue un avantage bien connu des CMOS; au repos, par exemple, elle consomme le *millième* de la 1101, et en dynamique, 2 % de la 1101 (laquelle exige, de plus, une double tension d'alimentation).

4. LA VOLATILITÉ DES RAM

On dit d'une mémoire qu'elle est *volatile* lorsque la coupure, accidentelle ou non, de l'alimentation, lui fait perdre l'information. C'est le cas général des RAM actuelles. Pour éviter un tel accident, on branche des batteries en tampon aux bornes des mémoires; encore faudra-t-il que ces dernières consomment le moins possible pour en réduire le volume.

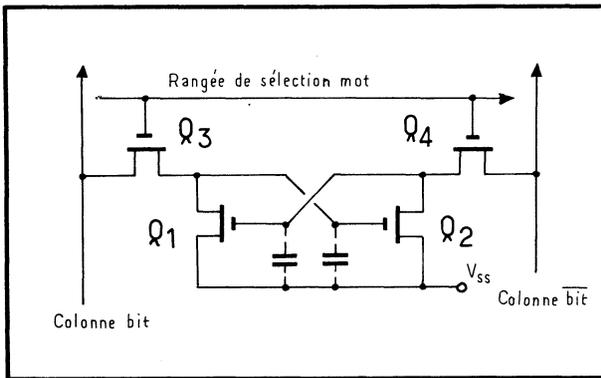


Fig. 9. — Cellule-type semi-dynamique à 4 MOS.

A nouveau se manifeste ici l'avantage des CMOS : de petites batteries garantissent la conservation des données enregistrées même pendant de longues heures. Dans le cas d'un système mémoire de faible importance, deux petites piles de 1,5 V en série (donc 3 V), pourront même se révéler suffisantes si on les branche en paral-

lèle sur les mémoires; elles relaieront l'alimentation défailante si l'on a prévu le dispositif automatique de mise en service.

5. RAM PSEUDO-DYNAMIQUES

La langue n'étant pas figée, on appelle indifféremment *pseudo-statiques*, *pseudo-dynamiques*, *demi* ou *semi-statiques* ou *dynamiques* des RAM qui, tout en conservant une structure de base rappelant un flip-flop incomplet, n'en mémorisent pas moins l'information sur leurs capacités intrinsèques. Avec les RAM réellement dynamiques, la pseudo-structure de flip-flop disparaît. L'avantage de la formule pseudo-dynamique est double :

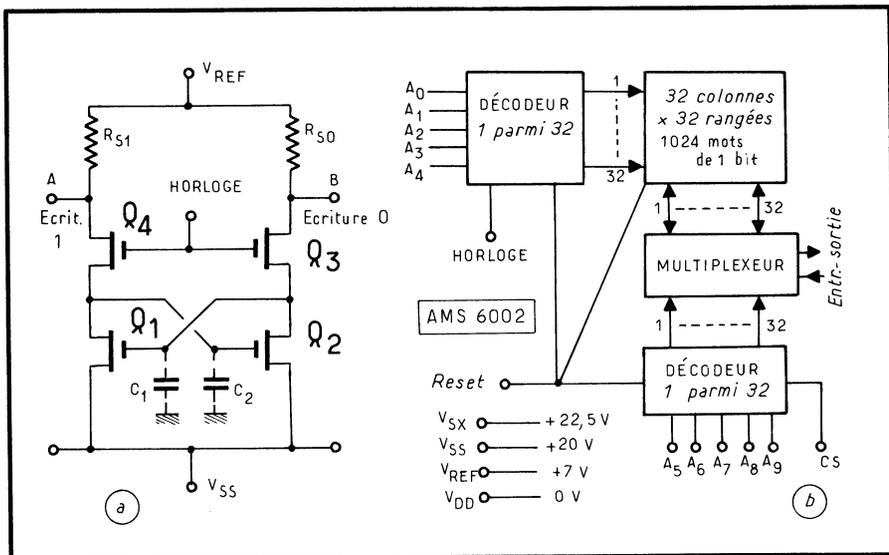


Fig. 10. — La RAM pseudo-dynamique type AMS 6002.

1. D'une part, elle se contente de 4 MOS par cellule au lieu de 6 pour le statique (mais de 3 pour le dynamique), ce qui réduit la surface occupée.

2. D'autre part, elle obéit à des règles d'utilisation moins astreignantes que celles qu'imposent les mémoires dynamiques.

Le cellule semi-dynamique à 4 MOS affecte la structure d'un flip-flop, mais sans charges (fig. 9); de ce fait, l'information doit être stockée dans les capacités de structure de Q_1 et Q_2 . Au départ d'un cycle, les colonnes *bit* sont préchargées à un potentiel de l'ordre de V_{DD} . La rangée de sélection (*mot*) est autorisée et Q_3 et Q_4 sont conducteurs. Si le signal au gate Q_1 est au niveau *haut*, Q_1 conduit et décharge la colonne *bit* à un potentiel proche de V_{SS} . En même temps, le gate de Q_1 est « rafraîchi » au potentiel de la colonne *non-bit* car Q_2 reste bloqué, tandis

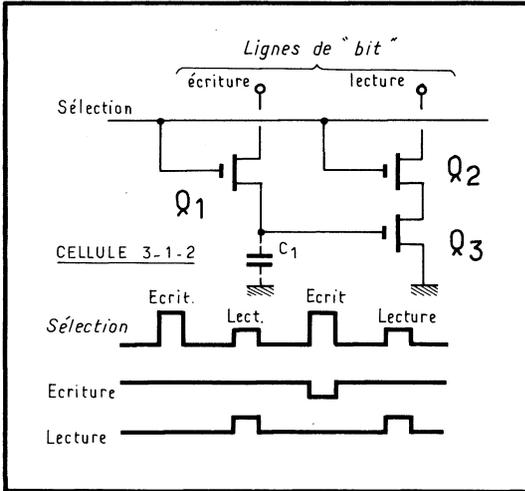


Fig. 11. — Cellule 3-1-2

que le gate de Q₂ est ramené à V_{SS}, potentiel de la colonne *bit*. Après quoi, on peut lire via Q₃ et Q₄, lorsque la rangée *mot* est activée, l'état du flip-flop. Pour écrire, les informations complémentaires sont introduites sur les colonnes *bit*. Ainsi, la procédure de rafraîchissement est identique à celle d'accès, ce qui explique que la lecture provoque un *rafraîchissement* automatique de la mémoire.

L'une des premières mémoires à appliquer ce principe a été la 6002 de *Advanced Memory Systems* (AMS); mais il a surtout été popularisé par la 3534 de *Fairchild*, dont la cellule est très exactement celle dessinée. Il n'y a pas besoin ici de commande de précharge, et les recouvrements des chronogrammes sont éliminés; la synchronisation, la commande et les interfaces sont, par conséquent, simplifiés. Cette RAM de 1 024 bits consomme 2 mW max. au repos, pour un accès de 150 à 300 ns max.,

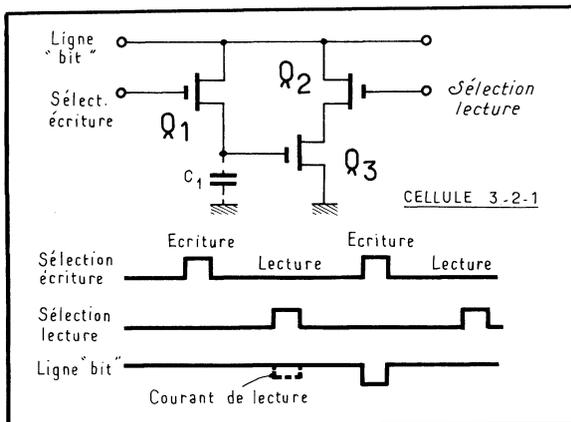


Fig. 12. — Cellule 3-2-1.

et un cycle de 250 à 480 ns max. En raison de la structure semi-dynamique, il faut procéder à un rafraîchissement systématique des cellules toutes les 2 ms.

Quant à la 6002 réalisée par AMS, sa structure fait apparaître une horloge (fig. 10) attaquant les gates de MOS, en porte AND avec ceux du flip-flop, la cellule comprenant 6 composants. Cette horloge sert de sélection de cellule (adresse) puisqu'elle l'active lorsqu'elle est portée à 1. L'écriture se fait alors normalement, la lecture mettant en évidence une tension différentielle entre A et B; le rafraîchissement s'effectue toutes les 2 ms, en 32 cycles.

Cette mémoire de 1 024 mots de 1 bit a un accès de 150 ns maximum, et un cycle de 250 à 290 ns, car au temps d'accès s'ajoute le temps de précharge, obtenue lorsque la ligne « reset » est basse. Présenté en boîtier 22 broches, ce CI consomme 2 mW au repos et 180 mW en opération.

6. RAM DYNAMIQUES

Les RAM dynamiques constituent un chapitre important des mémoires *vives* :

- elles sont plus simples à fabriquer (moins de transistors par cellule : 3 ou 1);
- elles consomment peu ;
- elles permettent une grande densité d'intégration (meilleure ouverture vers les 16 K et au delà).

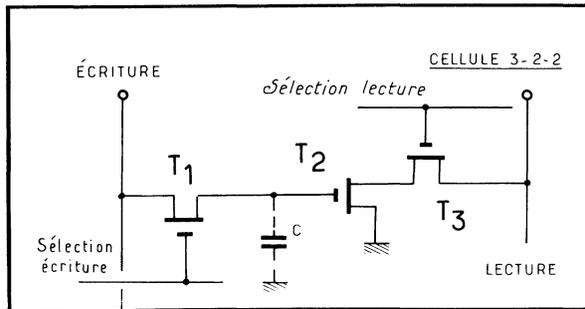


Fig. 13. — Cellule 3-2-2.

Les RAM dynamiques sont conçus à partir de cellules à trois ou à un seul transistor. Il existe 4 types de cellules dynamiques à 3 MOS. On les désigne souvent par un groupe de trois chiffres, et l'on trouve les 4 types suivants :

- a) La cellule 3-2-2 (3 transistors, 2 lignes d'adresse, 2 lignes de bit).
- b) La cellule 3-2-1 (3 transistors, 2 lignes d'adresse, 1 ligne de bit).
- c) La cellule 3-1-2 (3 transistors, 1 ligne d'adresse, 2 lignes de bit).
- d) La cellule 3-1-1 (3 transistors, 1 ligne d'adresse, 1 ligne de bit).

De telles cellules ont un temps d'accès de l'ordre de 300 ns et dissipent $10 \mu\text{W/bit}$ au repos, et 0,1 à 0,5 mW/bit en activité. Elles occupent entre 3 000 et 5 000 μm^2 de silicium, en canal P.

L'inscription de la cellule 3-1-2 (fig. 11) est assurée grâce à la commande *Sélection* qui débloque Q_1 et à l'application du niveau logique 0 ou 1 à l'entrée de la ligne *Bit* qui charge C_1 . Lors de la lecture, la ligne *Sélection* débloque Q_2 qui, avec Q_3 , constitue un AND dont l'état dépend de la charge de C_1 . La tension

de commande appliquée à *Sélection* doit être plus faible à la lecture qu'à l'écriture pour éviter une inscription erronée.

Pour tourner cette dernière difficulté typique de la 3-1-2, on utilise les 2 lignes d'adresse (sélection) de la cellule 3-2-1, où ne subsiste plus qu'une ligne commune de bit pour l'écriture et la lecture (fig. 12). Les ordres de sélection sont désormais d'amplitude égale. Dans la 3-2-2, on a rétabli les deux lignes de bit qui s'ajoutent aux deux lignes d'adresse, pour obtenir un fonctionnement très sûr et sans qu'aucune ligne ne serve une double fonction multiplexée (fig. 13). Si l'on multiplexe à la fois les lignes de lecture-écriture, et d'adresse, on aboutit à la cellule 3-1-1 (fig. 14), moins aisée d'emploi cependant car la lecture est destructive.

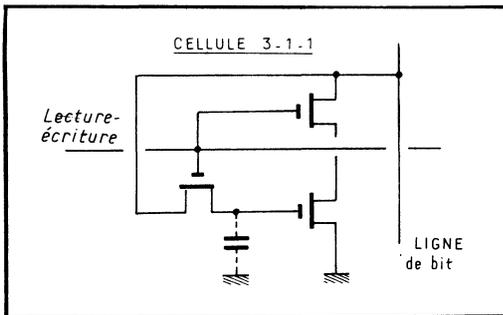


Fig. 14. — Cellule 3-1-1.

Si l'on mémorise l'information dans une capacité de structure, la plus simple des cellules est celle de la figure 15 où un seul transistor MOS apparaît. Ici, la lecture est destructive, et la charge stockée par C est redistribuée à la lecture aux capacités de la ligne *Bit*; de plus, la vitesse d'écriture est limitée par la capacité de C qui doit être suffisamment importante.

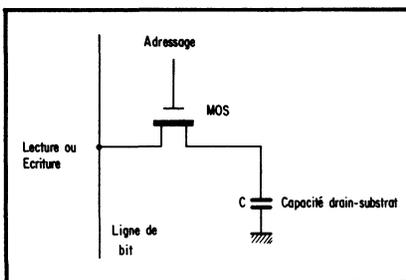


Fig. 15. — Cellule mémoire à un seul transistor MOS

Toutes les cellules dynamiques ici décrites présentent un trait commun : elles doivent être « rafraîchies » périodiquement, c'est-à-dire qu'il faut fournir aux capacités de structure en train de se décharger de nouvelles charges toutes les 2 ms environ. Enfin, il faut éventuellement compter aussi avec la « précharge » des lignes.

7. L'AVENIR DES RAM

Les marchés de l'informatique, qui constituent le principal client des mémoires, exigent des dispositifs toujours plus denses et plus rapides.

En 1978, les RAM dynamiques de 4 K étaient largement commercialisées et les 64 K étaient échantillonnées. Mais les technologies se remettant sans cesse en cause, on a vu apparaître, dès 1977, des RAM statiques plus rapides que leurs homologues dynamiques. Elles résultent de procédés plus fins comme de l'utilisation de nouvelles technologies, telle que la VMOS ou la HMOS. Les temps d'accès annoncés pouvaient être ramenés à quelque 30 à 50 ns, concurrençant ainsi les RAM bipolaires Schottky.

Il est certain que les progrès vont se poursuivre dans toutes les voies : densité, vitesse, consommation, facilité d'emploi, etc. Le progrès le plus décisif consisterait certainement à rendre non volatil ce type de mémoire, ce que l'on pense obtenir avec les MNOS (présentés dans le chapitre précédent).

En matière de densité, les CCD qui se prêtaient aux mémoires 64 K en 1978 devraient franchir allégrement la borne des 100 K, ainsi que les bulles magnétiques dont on attend d'ailleurs bien plus mais qui, dès l'origine, ne sont pas volatiles. Ces bulles magnétiques sont plus lentes que les CCD (accès en quelque 10 ms) mais devraient être bien moins chères, les CCD restant elles-mêmes plus économiques que les cellules à transistors (mais plus lentes également).

CHAPITRE XVI

LES MICROPROCESSEURS

L'une des plus belles applications des circuits intégrés, c'est certainement les microprocesseurs. Un microprocesseur est un circuit intégré LSI constituant l'unité centrale d'un ordinateur et par conséquent, commandé par un programme. On va en examiner ici les traits caractéristiques.

1. LOGIQUE CÂBLÉE ET LOGIQUE PROGRAMMÉE

La logique classique, pour résoudre ses problèmes, fait appel à des circuits câblés entre eux. La structure résultante est par conséquent « définitive ». C'est ce qu'on appelle la *logique câblée*.

Or, certains circuits exécutent des fonctions qui, au cours d'un unique processus, peuvent être appelées plusieurs fois. Au lieu d'en multiplier le nombre chaque fois qu'on en a besoin, on peut les faire intervenir *par programme*, à la demande; ainsi, un seul circuit d'un type donné est nécessaire (par exemple, un compteur, un OU exclusif, etc.).

Dans un circuit intégré complexe, LSI, on peut alors inclure toutes les fonctions de base utiles à la logique. A chaque instant, on s'arrangera pour ne mettre en service que le, ou les circuits indispensables, en bloquant les autres, ce qui s'obtient très facilement à l'aide de simples portes. Celles-ci seront commandées en séquences par un *programme*, suite d'instructions qui mettront successivement en activité les fonctions concernées. C'est ce qu'on appelle la *logique programmée*.

Par exemple, l'ensemble des fonctions arithmétiques et logiques sera exécutée par une *unité arithmétique et logique*, appelée ALU de l'américain « *arithmetic and logic unit* » contenant des portes logiques, un additionneur, un décaleur, un inverseur..., activés à la demande sur appel de l'instruction correspondante. Ainsi, au lieu de multiplier les circuits, on multiplie les instructions.

Les circuits, c'est le « hardware », francisé en *matériel*, alors que les instructions, c'est du « software » devenu le *logiciel*. En programmé, on remplace donc le matériel du *câblé* par du *logiciel*, plus rapidement développé, moins cher, plus fiable (des instructions ne tombent pas en panne!), moins volumineux et de consommation... disons quasiment nulle. Le seul inconvénient du programmé est sa vitesse d'exécution moindre par rapport au câblé, à technologies identiques.

2. STRUCTURE D'UN MICROPROCESSEUR

Outre l'unité arithmétique et logique, le microprocesseur doit posséder des circuits capables de recevoir les instructions (un mot binaire), de les décoder et de les faire exécuter, le tout synchronisé par une horloge. Il doit également disposer de quelques registres, qui serviront à stocker temporairement des informations utiles en cours de fonctionnement. De ce fait, la structure de principe d'un microprocesseur comporte les quatre sections de la figure 1, agissant de la façon suivante :

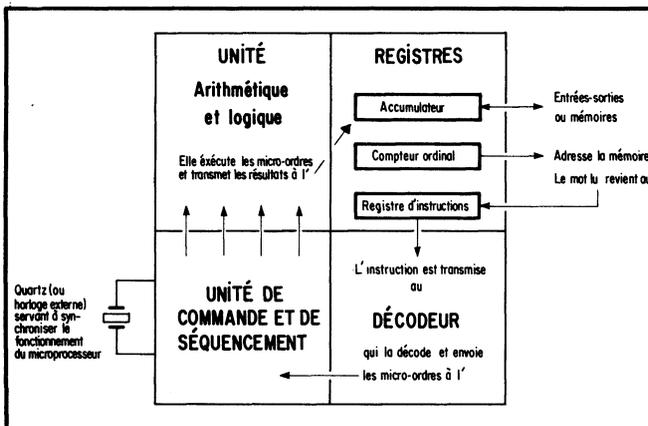


Fig. 1. — Structure (et fonctions) d'un microprocesseur : c'est un circuit intégré unique, LSI, commandé par un programme.

- *Un compteur ordinal* stocke, sous forme binaire, une « adresse ». Celle-ci désigne l'une des cellules d'une mémoire externe, RAM, ROM, PROM, EPROM... Cette adresse est envoyée à la mémoire qui, en retour, expédie son contenu vers le *registre d'instructions*.

- Le *registre d'instructions* transmet son information à un *décodeur* qui la traduira par une séquence de micro-ordres, fournis à l'*unité de commande et de cadencement*.

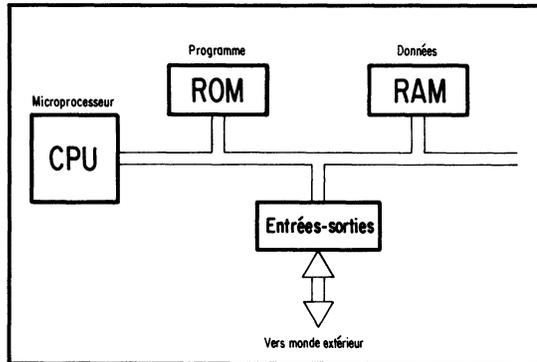
- Assujettie à une horloge (pilotée par un quartz externe, par exemple), l'*unité de commande et de cadencement* commande l'exécution des micro-ordres par ouverture ou fermeture de portes. Par exemple, elle mettra en service à un moment donné l'*additionneur* contenu dans l'ALU. Le résultat de l'opération sera logé dans un autre registre, appelé *accumulateur*.

- L'*accumulateur* servira ainsi à stocker des données ou des résultats de traitement, avant que ceux-ci ne soient relogés dans une cellule d'une mémoire externe, ou envoyés vers une sortie.

Les données sont constituées par des *mots*, ensemble de bits (binaires) traités en parallèle. Le nombre de bits en parallèle qu'accepte un microprocesseur sert souvent à un premier classement : on dira d'un microprocesseur qu'il traite des mots de un quartet (4 bits), de 1 octet (8 bits), de 12 bits ou de 16 bits...

Cet ensemble constitutif du microprocesseur représente très précisément l'unité centrale d'un ordinateur, ce que l'on appelle CPU, de « *central processing unit* ». Ce terme CPU est ainsi synonyme de microprocesseur, mais on utilise encore MPU, pour « *micro processing unit* ».

Fig. 2. — Organisation de base d'un micro-ordinateur.



Pour pouvoir être exploité, ce CPU doit être connecté à des mémoires qui stockent le programme et les données, et à des circuits d'entrées-sorties qui le relient au monde extérieur. L'ensemble de ces circuits, auxquels s'ajoutent une horloge, une alimentation, et éventuellement d'autres éléments (circuits d'initialisation, commande de bus, décodeurs...) constitue le *micro-ordinateur*. Rappelons qu'un bus est une voie de liaison parallèle (pour transmettre des mots de n bits en parallèle) desservant plusieurs éléments.

Parce qu'en général, le programme exécuté par le micro-ordinateur est immuable, on peut le loger dans une ou plusieurs mémoires mortes (ROM, PROM ou EPROM). Par contre, les données étant essentiellement des variables, on les confiera à des RAM. De ce fait, le synoptique le plus simple d'un système comprend (fig. 2) le CPU, une ROM, une RAM, et un circuit d'entrées-sorties reliant le micro à des périphériques (télé-imprimeur, clavier, écran, unité de mémoire de masse, à disques souples par exemple, etc.) ou à des dispositifs industriels (capteurs, actionneurs, relais, etc.).

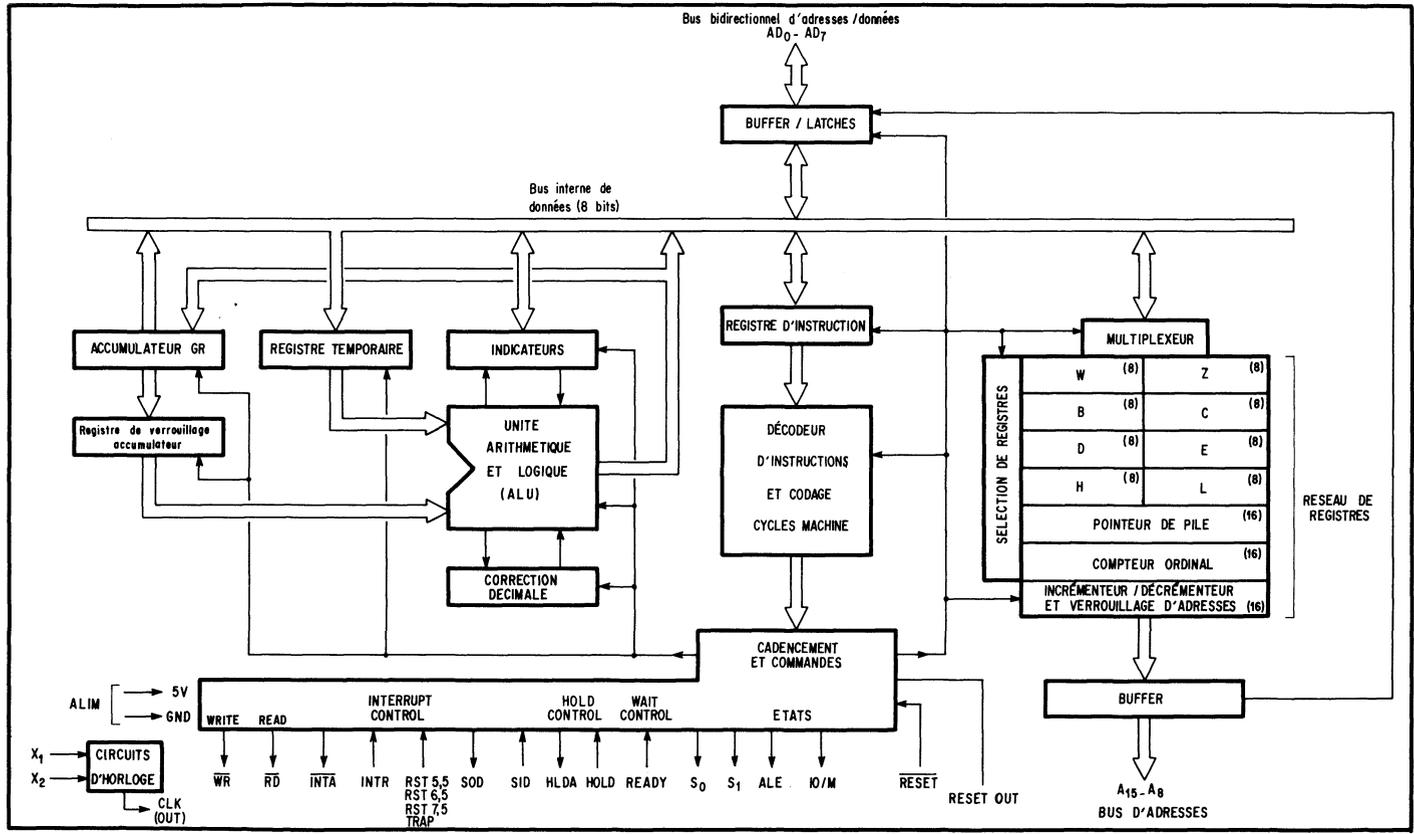


Fig. 3. — Synoptique du microprocesseur 8085.

3. L'EXEMPLE D'UN MICROPROCESSEUR : LE 8085

Il existe plusieurs dizaines de types de microprocesseurs, tant MOS que bipolaires. On considère que c'est la société *Intel* qui, la première, a commercialisé un microprocesseur, le 4004, vers 1970. Ce CPU en MOS canal P, de 4 bits, a été suivi par le premier 8 bits, appelé 8008. Ce dernier a été amélioré et transposé en NMOS pour devenir le 8080, dont le successeur, le 8085, a été commercialisé en 1977.

Le synoptique du 8085 est donné figure 3. On y reconnaîtra les éléments cités ci-dessus, mais en plus, on trouvera : des buffers, étages tampons pour les bus d'adresses et de données; une pile de registres dont ceux désignés par les lettres B, C, D, E, H et L mis à la disposition de l'utilisateur; deux registres W et Z réservés aux traitements internes; un pointeur de pile indiquant le niveau de remplissage d'une pile logicielle; un registre d'indicateurs d'états caractérisant certaines situations; un élément de correction décimale et des registres annexes. Développer plus avant leur fonctionnement et leur rôle dépasse le cadre de cet ouvrage.

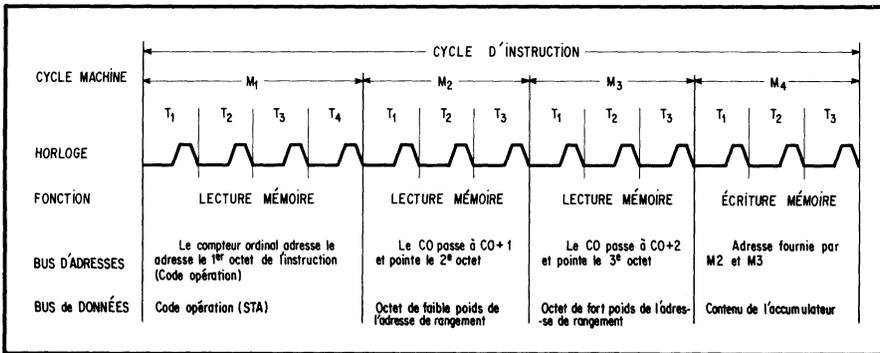


Fig. 4. — Chronogramme d'exécution d'une instruction de rangement en mémoire.

Alimenté sous +5 V, ce microprocesseur dispose de circuits d'horloge interne, qu'il faut piloter à l'aide d'une référence de temps qui peut être un quartz à 6 MHz. L'essentiel des opérations qu'il exécute peut se classer en lectures ou écritures; une écriture ou une lecture est exécutée en 1 cycle machine, qui dure plusieurs cycles d'horloge : ces derniers sont appelés *microcycles*. Une instruction est exécutée en une séquence de 1 à 5 cycles machine, chacun de ceux-ci comprenant de 3 à 6 microcycles.

Ainsi, soit une instruction disant de ranger le contenu de l'accumulateur dans une cellule mémoire dont l'adresse est fournie par cette même instruction. Cette instruction sera codée en binaire, mais pour plus de facilité, on passe par un code mnémotechnique qui sera, ici : STA 384A, par exemple, pour : STA, « store accumulator » (ranger l'accumulateur) à l'adresse 384A (en hexadécimal). L'exécution de cette instruction est développée figure 4 :

- *M1, premier cycle machine* : l'adresse du compteur ordinal est émise sur le bus d'adresses pour aller chercher l'instruction STA logée en mémoire. Le cycle est appelé « fetch », ou recherche. Pendant le 4^e microcycle de M1, l'instruction STA est décodée et le CPU apprend qu'il doit continuer à lire la mémoire pour

trouver l'adresse 384A. Le compteur ordinal est incrémenté par le CPU qui lui ajoute 1.

• *M2, second cycle machine* : le compteur ordinal pointe désormais la cellule mémoire suivante, qui est lue à son tour. Mais celle-ci ne contient qu'un mot de 1 octet, soit 2 digits hexadécimaux; or, on a demandé une adresse sur 4 digits hexadécimaux. Par conséquent, il faudra une autre lecture en mémoire. A la fin de M2, le compteur ordinal est encore incrémenté.

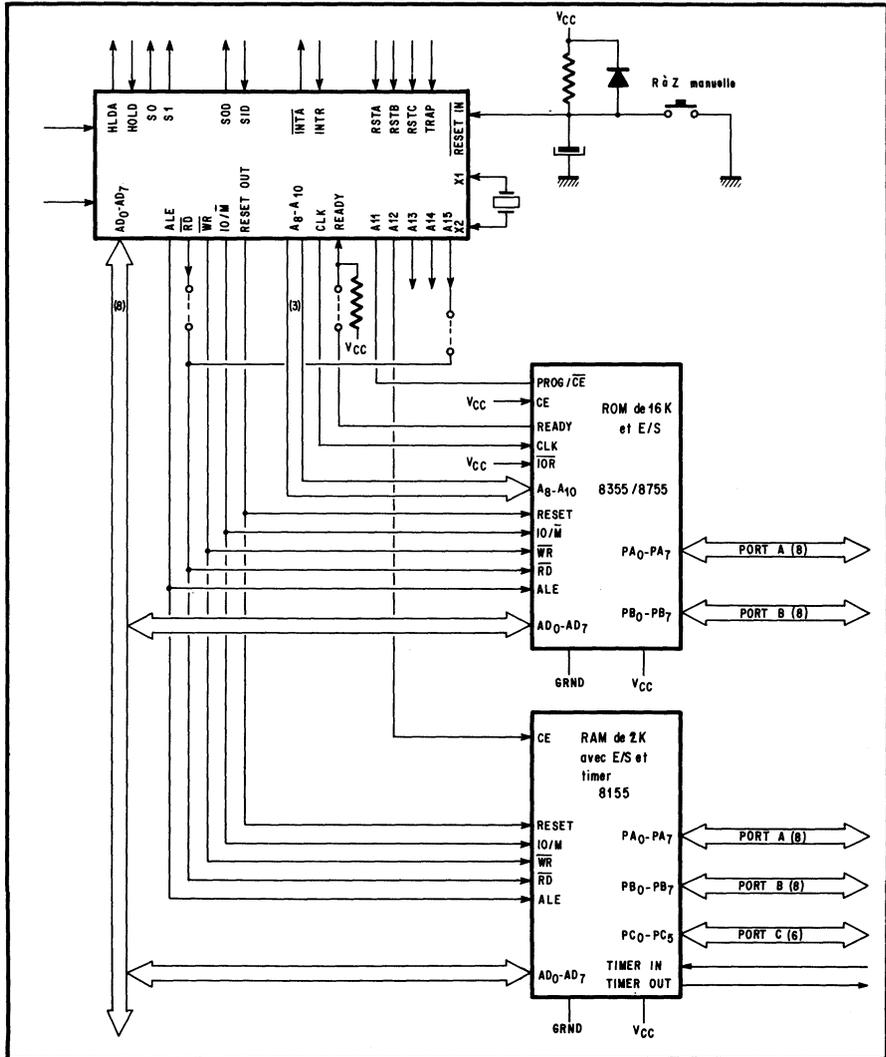


Fig. 5. — Micro-ordinateur de base construit autour d'un 8085. On peut accroître le nombre des circuits mémoires et d'entrées-sorties, en particulier. Il faut aussi y ajouter une alimentation.

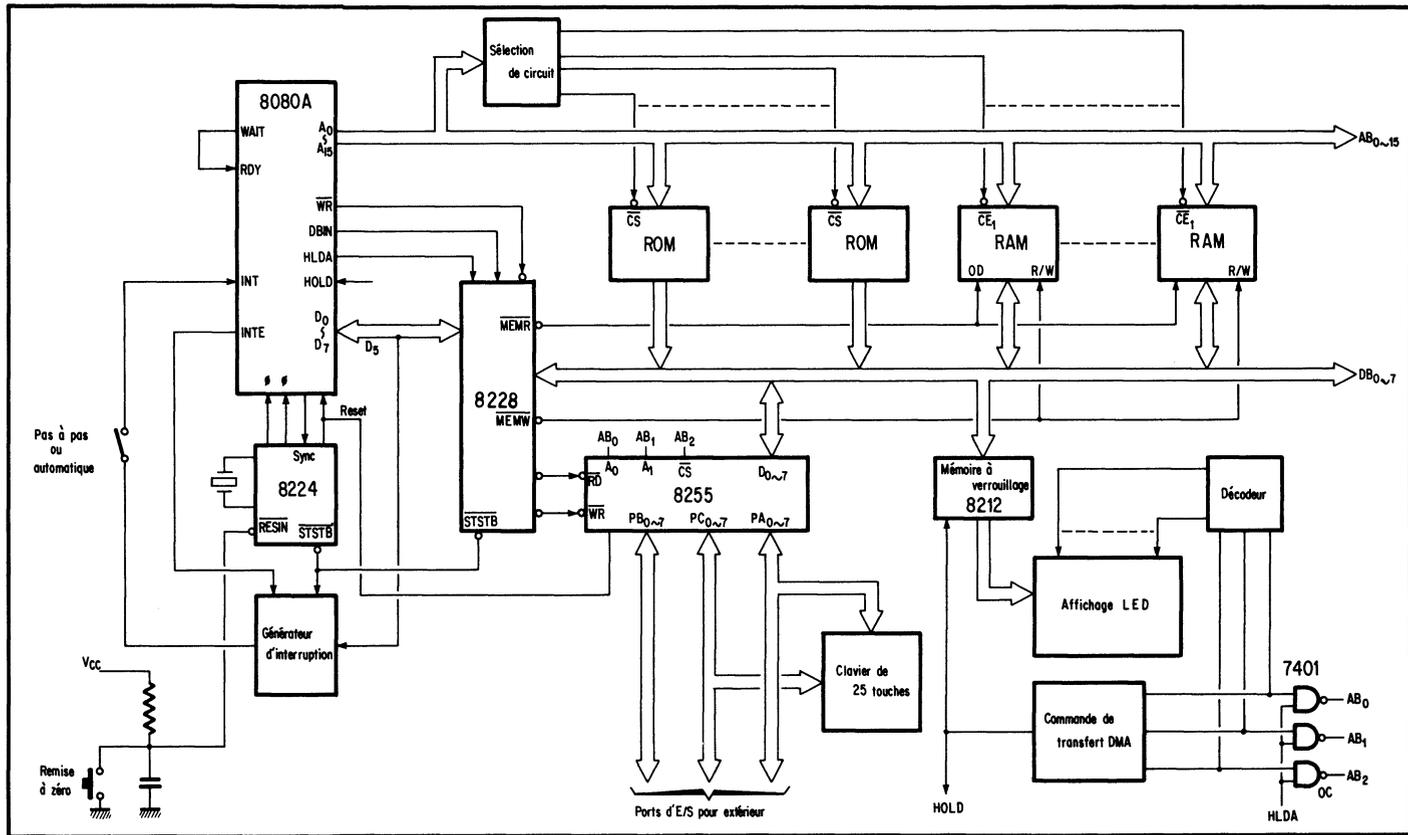


Fig. 6. — Exemple d'un micro-ordinateur complet conçu pour l'enseignement de la programmation : le MTS de ICS, ou le TK 80 de NEC.

- *M3, troisième cycle machine* : on recommence la même opération afin de disposer de l'adresse complète 384A.
- *M4, quatrième cycle* : il reste alors à émettre cette adresse 384A sur le bus d'adresses, à connecter l'accumulateur au bus de données et, au cours de ce dernier cycle d'écriture, à exécuter l'instruction prescrite.
- Après quoi on passera à l'instruction suivante.

Un cycle machine suffit, par exemple, pour exécuter une addition, à partir du moment où le cumulateur et le registre se trouvent déjà dans les registres du CPU. Ainsi, l'addition demande 1,3 μ s.

Comme tous les microprocesseurs, le 8085 dispose de plusieurs dizaines d'instructions possibles, dûment codées et imposées par le fabricant : *rangement, chargement, addition, soustraction, branchements conditionnels ou inconditionnels, etc.* La façon de les aligner pour établir un programme, le langage utilisé pour sa rédaction, les programmes annexes de soutien..., tout cela constitue le *logiciel* dont l'étude ne sera pas entreprise ici. On se souviendra simplement qu'*un programme n'est pas autre chose qu'un plan de travail*, et qu'il ne peut être rédigé qu'après qu'un organigramme détaillé des fonctions à exécuter ait été dressé.

Sur la logique câblée, le programme offre encore l'avantage du pouvoir être aisément modifié, en période d'étude, de mise au point, de réalisation, et même après, sur le terrain : on peut toujours modifier les instructions, à la limite en changeant quelques circuits ROM ou PROM, ou en ré-inscrivant les EPROM, sur le terrain et même plusieurs années après la mise en service d'un équipement (pour ajouter de nouvelles fonctions, par exemple), ce qui est difficilement concevable en logique câblée.

Avec ses mémoires RAM et ROM, associées à des circuits d'entrées-sorties, le 8085 mène à un système micro-ordinateur dont la version la plus dépouillée ne comporte que 3 circuits intégrés (fig. 5).

4. EXEMPLE D'UN MICRO-ORDINATEUR D'ENSEIGNEMENT

Si le MTS (« *microprocessor training system* » : *système d'apprentissage au microprocesseur*) est pris ici comme exemple, c'est parce qu'il constitue un micro-ordinateur complet type sur une seule carte imprimée de 18 \times 31 cm : le microprocesseur, un 8080 (très proche du 8085, mais moins évolué), des mémoires, des interfaces, etc., ainsi qu'un clavier hexadécimal avec touches de fonctions qui sert à introduire les programmes et les données, et un affichage sur 8 digits hexadécimaux. Son synoptique est donné figure 6.

Il dispose de 1 K octet de mémoires EPROM, programmées par le concepteur (le groupe américain ICS qui dispense des cours de micro-informatique et autres) et contenant un programme *moniteur*, espèce de chef d'orchestre qui gère le fonctionnement de l'ensemble, scrute le clavier, commande l'affichage, etc.

Le circuit 8224 est un circuit d'horloge, le 8080 ne comprenant pas ceux-ci. Le 8228 constitue un tampon entre le CPU et les bus. Le 8255 est le circuit d'entrées-sorties. L'affichage sur diodes électroluminescentes demande quelques circuits annexes.

Le système est alimenté par deux tensions, +5 et +12 V. Les mémoires RAM étant des DS101, de 256 mots de 4 bits, en technologie CMOS, elles peuvent

être rendues non volatiles par l'adjonction de deux petites piles 1,5 V en série, soit 3 V, qui, branchées aux bornes des RAM lorsque l'alimentation est coupée, leur conservent les informations quasi indéfiniment.

De nombreux systèmes du type de ce MTS ont été commercialisés. Relativement économiques, ils permettront s'ils sont accompagnés de bons manuels, d'apprendre la programmation en se jouant, la réalisation d'un micro-ordinateur (matériel) n'offrant par ailleurs aucune difficulté aux électroniciens.

5. LES TROIS CATÉGORIES DE MICROPROCESSEURS

En pratique, l'utilisateur constate que le marché lui offre trois grandes catégories de microprocesseurs :

A) *Les microprocesseurs proprement dits* : ce sont les éléments qu'on vient de définir. Ils constituent donc le CPU, ou unité centrale de traitement, de la future machine informatique ou du système à réaliser.

Ainsi, la figure 7 montre comment est constitué un système avec un 6800, de *Sescomem* ou *Motorola* (ou *AMI...*).

B) *Les microtranches*, dont le type est fourni par le 2901 de *AMD* (ou *Sescomem*). On n'insistera pas ici sur cette catégorie, car elle relève essentiellement des technologies bipolaires, TTL et ECL, qui n'offrent pas les mêmes facilités d'in-

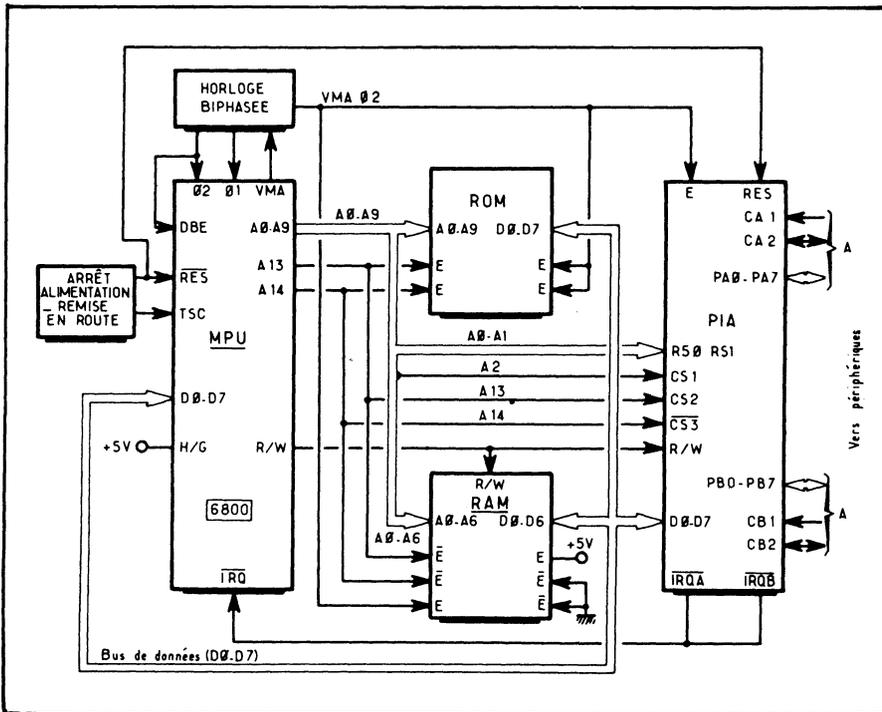


Fig. 7. — Système de base à 6800.

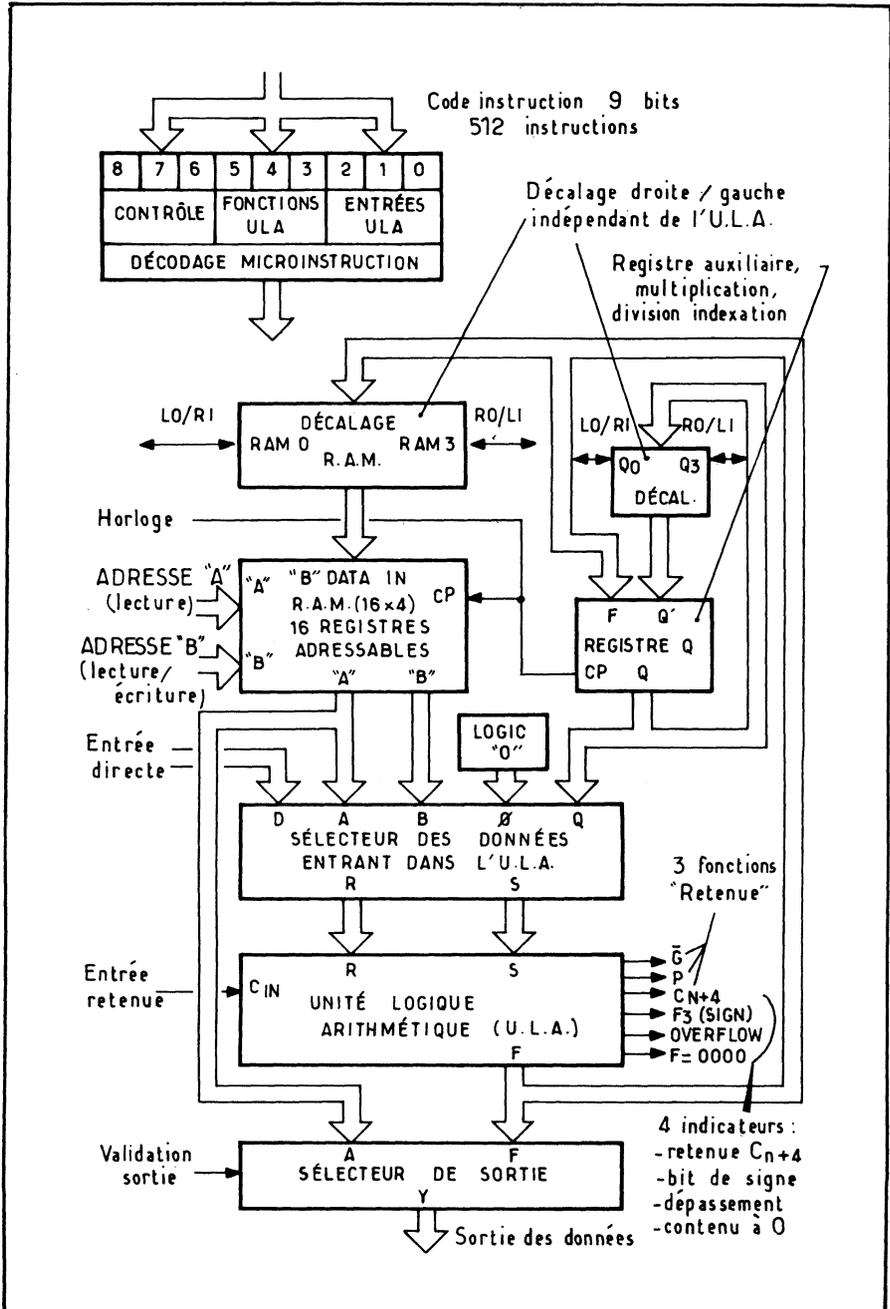


Fig. 8. — Composition du 2901, bipolaire Schottky, qui mène à la réalisation d'une unité arithmétique et logique « en tranches ».

tégration que les MOS. Par conséquent, l'intégration d'un microprocesseur complet sur une puce ne peut se faire et l'on recourt à de nombreux circuits, ou *tranches*, que l'on juxtapose.

L'une de ces tranches, celle du 2901, est dessinée figure 8; elle mène au système complet de la figure 9. Sa complexité est supérieure, prix qu'il faut payer pour une vitesse accrue. Avec le 2901, chaque tranche est de 4 bits : deux tranches mèneront à un système 8 bits, quatre à 16 bits, etc.

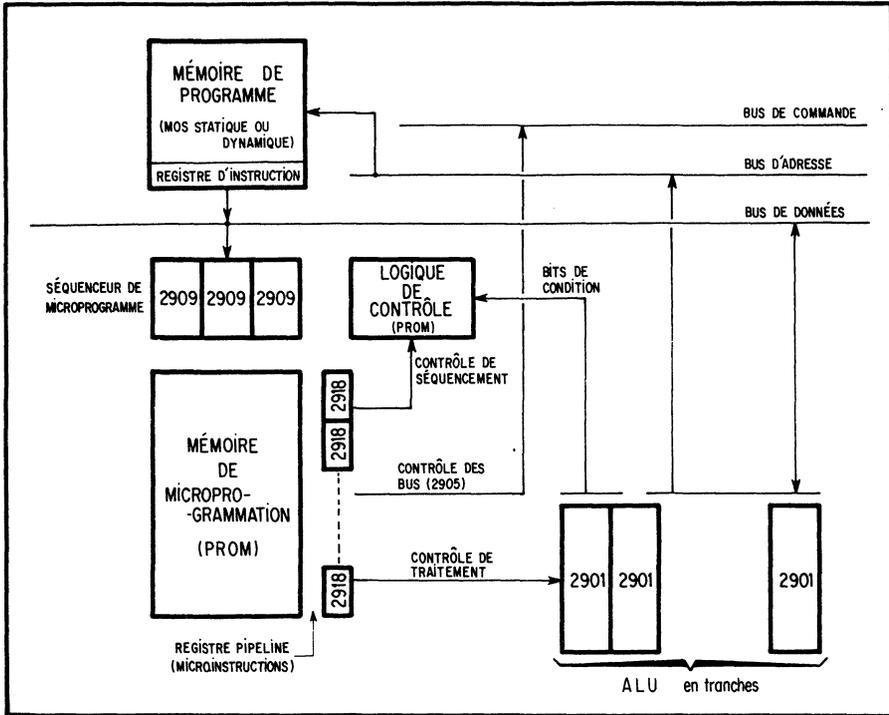


Fig. 9. — Système en tranche réalisé avec le 2901. Il regroupe de très nombreux circuits.

C) *Les microsystèmes.* Puisqu'un système à microprocesseur demande, outre le CPU, des mémoires ROM et RAM, des entrées-sorties, une horloge, etc..., ne pourrait-on regrouper tout cela sur une seule puce de silicium? C'est ce qui a été fait avec les microsystèmes.

Un exemple en est donné figure 10, c'est celui du TMS 1000 de Texas. Son synoptique montre que ce microsystème à mot de 4 bits regroupe une ROM de 1 024 octets, une RAM de 64 quartets, des entrées-sorties. Ajoutons que ce TMS 1000 existe en versions PMOS, NMOS ou CMOS.

D'autres microsystèmes existent avec mémoire morte incluse, ou alors avec EPROM (effaçable à l'ultraviolet).

En fait, ce sont des systèmes tels que le F8 de Fairchild qui ont ouvert la voie aux microsystèmes. Ce F8 permet, en effet, la réalisation d'un système complet avec deux circuits de base, le CPU et le PSU, dont le synoptique est illustré figure 11.

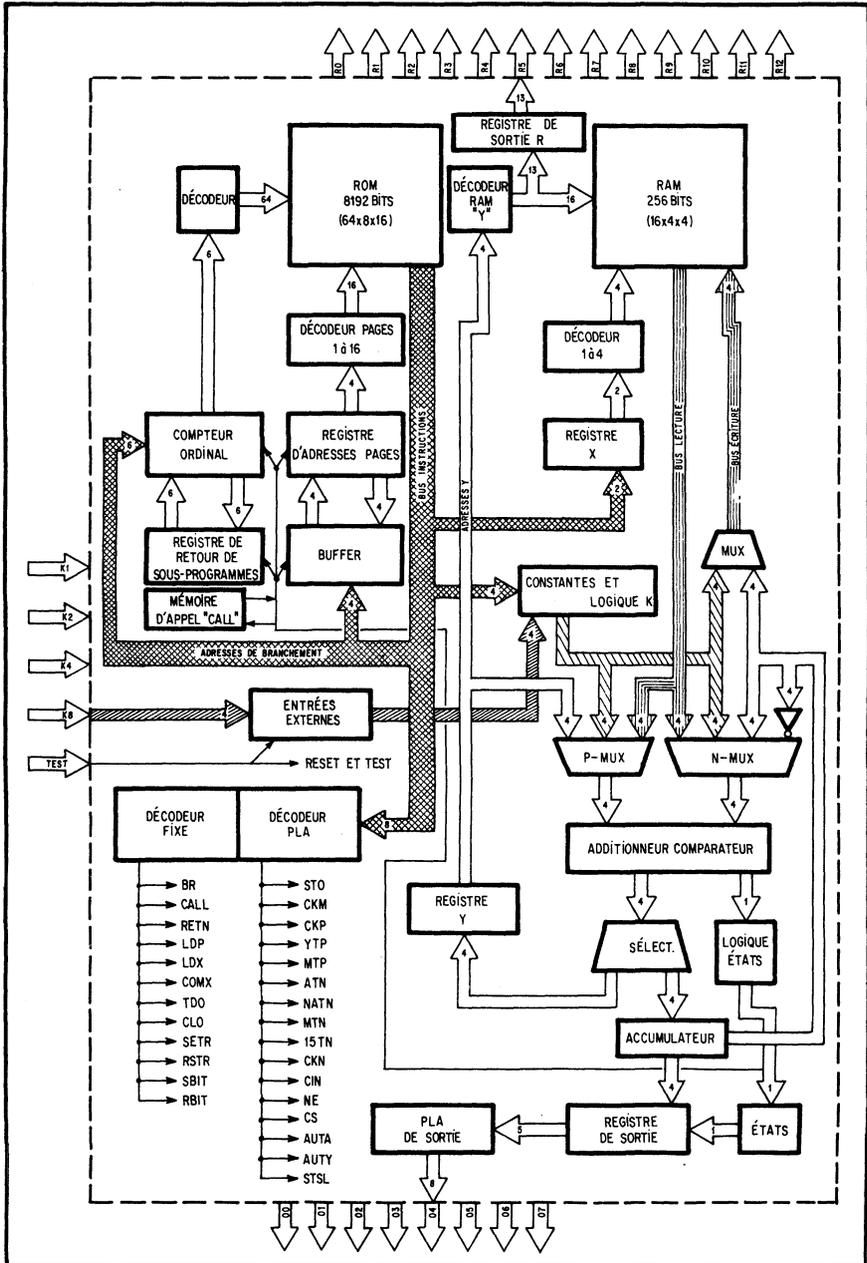


Fig. 10. — Organisation du TMS 1000, microsystème avec mémoires et entrées-sorties incorporées.

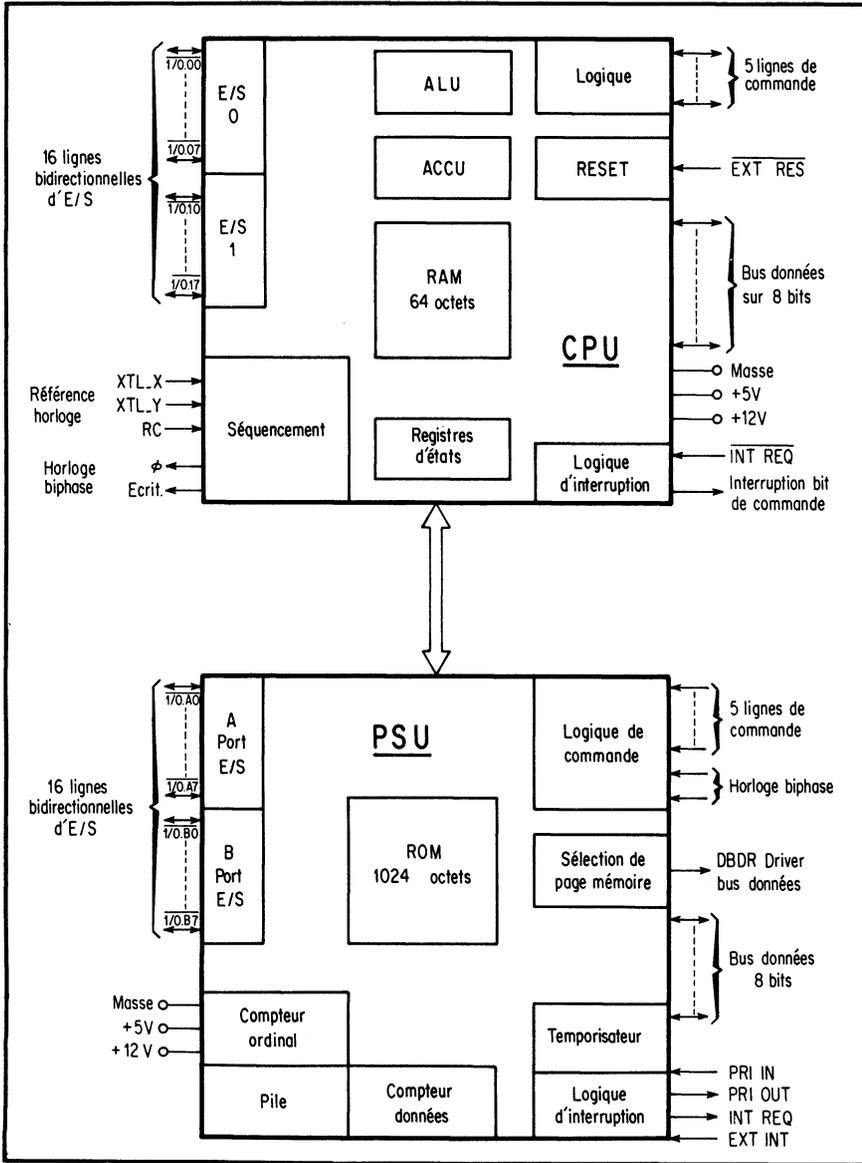


Fig. 11. — Système minimal F8 : il comprend deux circuits, le CPU avec RAM incorporée, et le PSU (pour « program storage unit ») avec ROM. C'est un système extrêmement simple, qui pourrait d'ailleurs même faire l'objet d'une réalisation en un unique circuit. Le F8 est proposé par Fairchild, son créateur, et par Mostek en seconde source, en particulier.

6. LE LOGICIEL

L'aspect le plus important du système à microprocesseur, tout au moins celui qui dérouterait le plus les électroniciens, de prime abord, c'est le logiciel. Par *logiciel*, qu'on oppose à *matériel*, on entend tout ce qui est programmes, modes d'emploi, et même documentations explicatives.

Qu'est-ce qu'un programme? C'est le plan de travail du microprocesseur. Un programme se compose d'une série d'*instructions*, alignées dans l'ordre, les unes en dessous des autres.

Ces instructions sont, dans la majorité des cas et actuellement, rédigées dans ce qu'on appelle le *langage d'assemblage* : les ordres élémentaires sont codés à l'aide de symboles mnémoniques, car ces derniers sont plus faciles à retenir. Le tableau donne ainsi, à titre d'exemple, le jeu d'instructions complet proposé par *Intel* pour le 8085; en effet, c'est le fabricant qui, ayant conféré à son microprocesseur les possibilités qu'il a jugées préférables, en dresse la liste. L'utilisateur devra

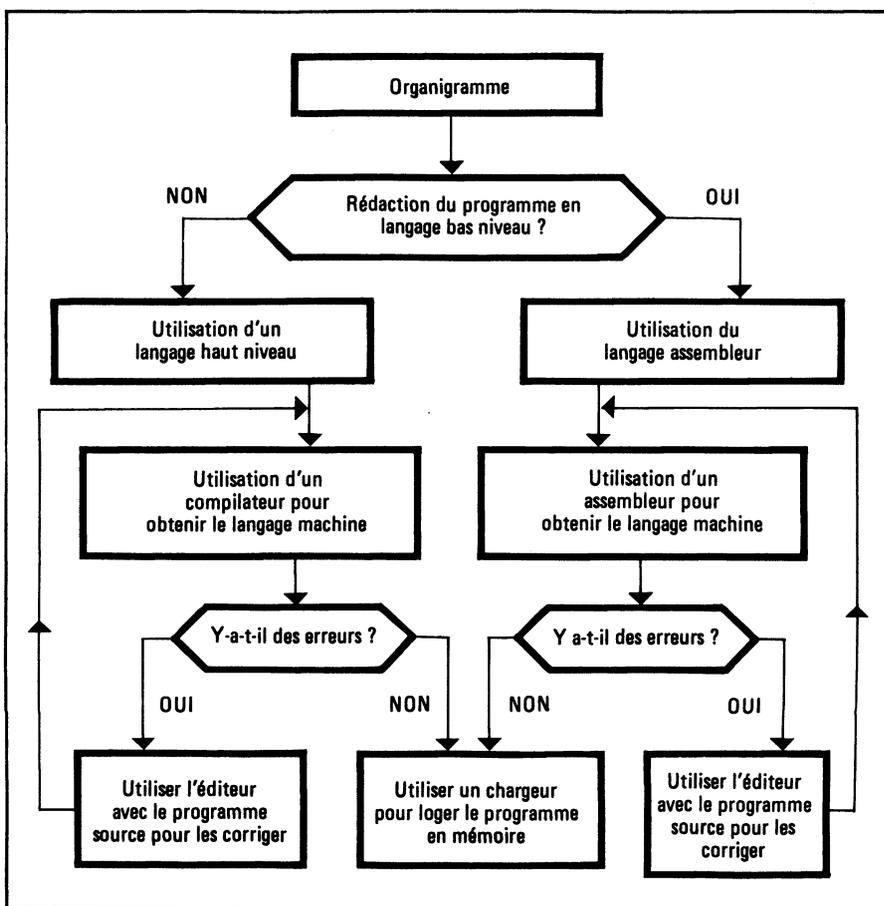


Fig. 12. — Les principes de développement d'un programme d'application.

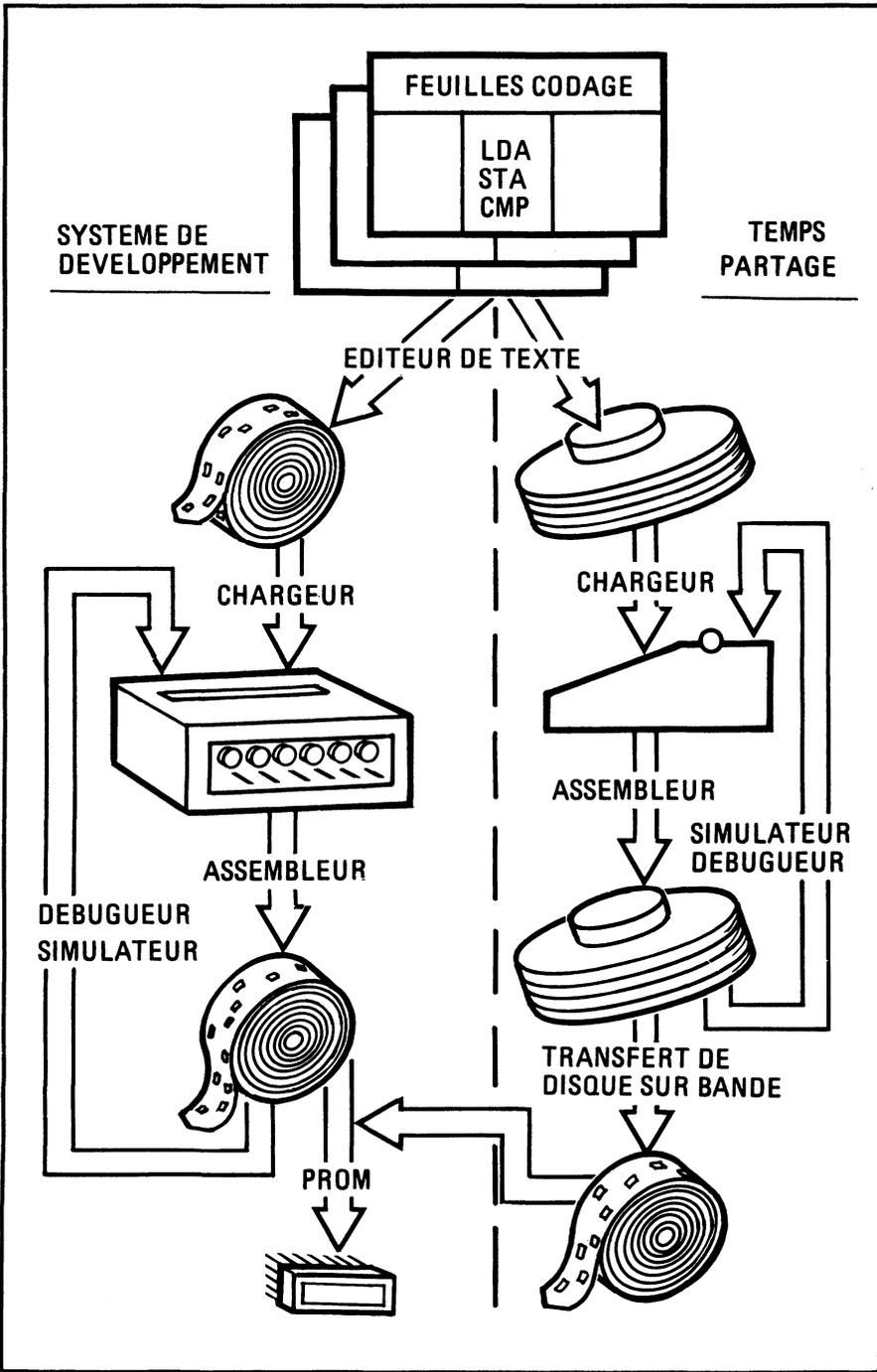


Fig. 13. — Deux façons de réaliser un programme, selon que l'on passe par un système de développement ou par des cross-logiciels en temps partagé.

Le jeu d'instructions du 8085

Mnemonic	Description	Instruction Code ^[1]								Clock ^[2] Cycles	Mnemonic	Description	Instruction Code ^[1]								Clock ^[2] Cycles
		D ₇	D ₆	D ₅	D ₄	D ₃	D ₂	D ₁	D ₀				D ₇	D ₆	D ₅	D ₄	D ₃	D ₂	D ₁	D ₀	
MOV <i>r₁, r₂</i>	Move register to register	0	1	0	0	0	S	S	S	4	RZ	Return on zero	1	1	0	0	1	0	0	0	6/12
MOV <i>M, r</i>	Move register to memory	0	1	1	1	0	S	S	S	7	RNZ	Return on no zero	1	1	0	0	0	0	C	0	6/12
MOV <i>r, M</i>	Move memory to register	0	1	D	D	D	1	1	0	7	RP	Return on positive	1	1	1	1	0	0	0	0	6/12
HLT	Halt	0	1	1	1	0	1	1	0	5	RM	Return on minus	1	1	1	1	1	0	0	0	6/12
MVI <i>r</i>	Move immediate register	0	0	D	D	0	1	1	0	7	RPE	Return on parity even	1	1	1	0	1	0	0	0	6/12
MVI <i>M</i>	Move immediate memory	0	0	1	1	0	1	1	0	10	RPO	Return on parity odd	1	1	1	0	0	0	0	0	6/12
INR <i>r</i>	Increment register	0	0	D	D	D	1	0	0	4	RST	Restart	1	1	A	A	A	1	1	1	12
DCR <i>r</i>	Decrement register	0	0	D	D	D	1	0	1	4	IN	Input	1	1	0	1	1	0	1	1	10
INR <i>M</i>	Increment memory	0	0	1	1	0	1	0	0	10	OUT	Output	1	1	0	1	0	0	1	1	10
DCR <i>M</i>	Decrement memory	0	0	1	1	0	1	0	1	10	LXI <i>B</i>	Load immediate register Pair B & C	0	0	0	0	0	0	0	1	10
ADD <i>r</i>	Add register to A	1	0	0	0	0	S	S	S	4	LXI <i>D</i>	Load immediate register Pair D & E	0	0	0	1	0	0	0	1	10
ADC <i>r</i>	Add register to A with carry	1	0	0	0	1	S	S	S	4	LXI <i>H</i>	Load immediate register Pair H & L	0	0	1	0	0	0	0	1	10
SUB <i>r</i>	Subtract register from A	1	0	0	1	0	S	S	S	4	LXI <i>SP</i>	Load immediate stack pointer	0	0	1	1	0	0	0	1	10
SBB <i>r</i>	Subtract register from A with borrow	1	0	0	1	1	S	S	S	4	PUSH <i>B</i>	Push register Pair B & C on stack	1	1	0	0	0	1	0	1	12
ANA <i>r</i>	And register with A	1	0	1	0	0	S	S	S	4	PUSH <i>D</i>	Push register Pair D & E on stack	1	1	0	1	0	1	0	1	12
XRA <i>r</i>	Exclusive Or register with A	1	0	1	0	1	S	S	S	4	PUSH <i>H</i>	Push register Pair H & L on stack	1	1	1	0	0	1	0	1	12
ORA <i>r</i>	Or register with A	1	0	1	1	0	S	S	S	4	PUSH <i>PSW</i>	Push A and Flags on stack	1	1	1	1	0	1	0	1	12
CMP <i>r</i>	Compare register with A	1	0	1	1	1	S	S	S	4	POP <i>B</i>	Pop register pair B & C off stack	1	1	0	0	0	0	0	1	10
ADD <i>M</i>	Add memory to A	1	0	0	0	0	1	1	0	7	POP <i>D</i>	Pop register pair D & E off stack	1	1	0	1	0	0	0	1	10
ADC <i>M</i>	Add memory to A with carry	1	0	0	0	1	1	1	0	7	POP <i>H</i>	Pop register pair H & L off stack	1	1	1	0	0	0	0	1	10
SUB <i>M</i>	Subtract memory from A	1	0	0	1	0	1	1	0	7											
SBB <i>M</i>	Subtract memory from A with borrow	1	0	0	1	1	1	1	0	7											
ANA <i>M</i>	And memory with A	1	0	1	0	0	1	1	0	7											
XRA <i>M</i>	Exclusive Or memory with A	1	0	1	0	1	1	1	0	7											
ORA <i>M</i>	Or memory with A	1	0	1	1	0	1	1	0	7											
CMP <i>M</i>	Compare memory with A	1	0	1	1	1	1	1	0	7											
ADI	Add immediate to A	1	1	0	0	0	1	1	0	7											
ACI	Add immediate to A with carry	1	1	0	0	1	1	1	0	7											

SUI	Subtract immediate from A	1	1	0	1	0	1	1	0	7
SBI	Subtract immediate from A with borrow	1	1	0	1	1	1	1	0	7
ANI	And immediate with A	1	1	1	0	0	1	1	0	7
XRI	Exclusive Or immediate with A	1	1	1	0	1	1	1	0	7
ORI	Or immediate with A	1	1	1	1	0	1	1	0	7
CPI	Compare immediate with A	1	1	1	1	1	1	0	7	
RLC	Rotate A left	0	0	0	0	0	1	1	1	4
RRC	Rotate A right	0	0	0	0	1	1	1	1	4
RAL	Rotate A left through carry	0	0	0	1	0	1	1	1	4
RAR	Rotate A right through carry	0	0	0	1	1	1	1	1	4
JMP	Jump unconditional	1	1	0	0	0	0	1	1	10
JC	Jump on carry	1	1	0	1	1	0	1	0	7/10
JNC	Jump on no carry	1	1	0	1	0	0	1	0	7/10
JZ	Jump on zero	1	1	0	0	1	0	1	0	7/10
JNZ	Jump on no zero	1	1	0	0	0	0	1	0	7/10
JP	Jump on positive	1	1	1	1	0	0	1	0	7/10
JM	Jump on minus	1	1	1	1	1	0	1	0	7/10
JPE	Jump on parity even	1	1	1	0	1	0	1	0	7/10
JPO	Jump on parity odd	1	1	1	0	0	0	1	0	7/10
CALL	Call unconditional	1	1	0	0	1	1	0	1	18
CC	Call on carry	1	1	0	1	1	1	0	0	9/18
CNC	Call on no carry	1	1	0	1	0	1	0	0	9/18
CZ	Call on zero	1	1	0	0	1	1	0	0	9/18
CNZ	Call on no zero	1	1	0	0	0	1	0	0	9/18
CP	Call on positive	1	1	1	1	0	1	0	0	9/18
CM	Call on minus	1	1	1	1	1	1	0	0	9/18
CPE	Call on parity even	1	1	1	0	1	1	0	0	9/18
CPO	Call on parity odd	1	1	1	0	0	1	0	0	9/18
RET	Return	1	1	0	0	1	0	0	1	10
RC	Return on carry	1	1	0	1	1	0	0	0	6/12
RNC	Return on no carry	1	1	0	1	0	0	0	0	6/12

POP PSW	Pop A and Flags off stack	1	1	1	1	0	0	0	1	10
STA	Store A direct	0	0	1	1	0	0	1	0	13
LDA	Load A direct	0	0	1	1	1	0	1	0	13
XCHG	Exchange D & E, H & L Registers	1	1	1	0	1	0	1	1	4
XTHL	Exchange top of stack, H & L	1	1	1	0	0	0	1	1	16
SPHL	H & L to stack pointer	1	1	1	1	1	0	0	1	6
PCHL	H & L to program counter	1	1	1	0	1	0	0	1	6
DAD B	Add B & C to H & L	0	0	0	0	1	0	0	1	10
DAD D	Add D & E to H & L	0	0	0	1	1	0	0	1	10
DAD H	Add H & L to H & L	0	0	1	0	1	0	0	1	10
DAD SP	Add stack pointer to H & L	0	0	1	1	1	0	0	1	10
STAX B	Store A indirect	0	0	0	0	0	0	1	0	7
STAX D	Store A indirect	0	0	0	1	0	0	1	0	7
LDAX B	Load A indirect	0	0	0	0	1	0	1	0	7
LDAX D	Load A indirect	0	0	0	1	1	0	1	0	7
INX B	Increment B & C registers	0	0	0	0	0	0	1	1	6
INX D	Increment D & E registers	0	0	0	1	0	0	1	1	6
INX H	Increment H & L registers	0	0	1	0	0	0	1	1	6
INX SP	Increment stack pointer	0	0	1	1	0	0	1	1	6
DCX B	Decrement B & C	0	0	0	0	1	0	1	1	6
DCX D	Decrement D & E	0	0	0	1	1	0	1	1	6
DCX H	Decrement H & L	0	0	1	0	1	0	1	1	6
DCX SP	Decrement stack pointer	0	0	1	1	1	0	1	1	6
CMA	Complement A	0	0	1	0	1	1	1	1	4
STC	Set carry	0	0	1	1	0	1	1	1	4
CMC	Complement carry	0	0	1	1	1	1	1	1	4
DAA	Decimal adjust A	0	0	1	0	0	1	1	1	4
SHLD	Store H & L direct	0	0	1	0	0	0	1	0	16
LHLD	Load H & L direct	0	0	1	0	1	0	1	0	16
EI	Enable Interrupts	1	1	1	1	1	0	1	1	4
DI	Disable interrupt	1	1	1	1	0	0	1	1	4
NOP	No-operation	0	0	0	0	0	0	0	0	4
RIM	Read Interrupt Mask	0	0	1	0	0	0	0	0	4
SIM	Set Interrupt Mask	0	0	1	1	0	0	0	0	4

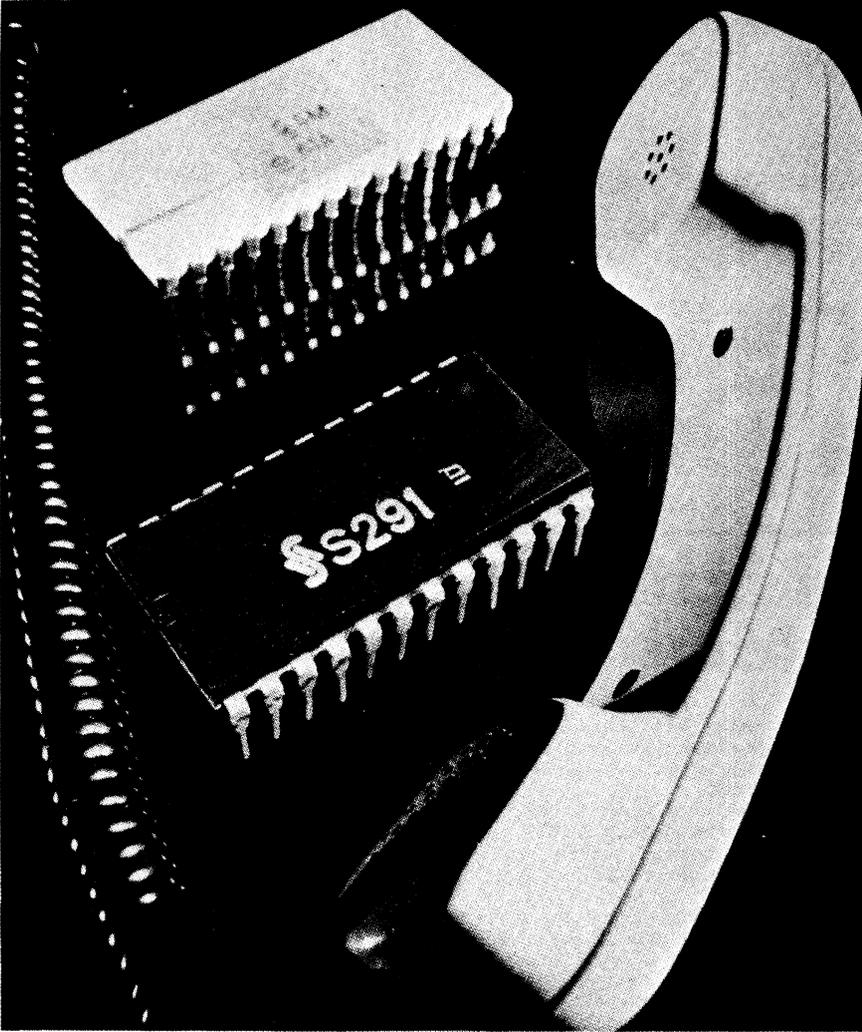
- NOTES: 1. DDD or SSS - 000 B - 001 C - 010 D - 011 E - 100 H - 101 L - 110 Memory - 111 A.
2. Two possible cycle times, (6/12) indicate instruction cycles dependent on condition flags.

ensuite jongler avec celles-ci. Dans ce tableau, conservé ici en anglais (mais jusqu'à présent, on n'a pas tenté de traduire en français les mnémoniques de base), on trouve en une première colonne ce mnémonique, puis la description succincte de la fonction exécutée, enfin le code correspondant en binaire car en définitive, les circuits intégrés logiques ne connaissent que cette langue, et enfin, la durée de l'exécution de l'instruction en cycles d'horloge. Il suffit de multiplier le nombre de ces microcycles par la période d'horloge pour obtenir le temps exact d'exécution.

Pour rédiger un programme, on procède d'abord à l'analyse des fonctions, on en trace un *ordinogramme* (ou organigramme), puis on aligne les instructions qui doivent intervenir à chaque étape. Cela fait, ces codes mnémoniques devront être traduits en binaire à l'aide d'un dictionnaire de traduction, le plus souvent sous forme de programme spécialisé appelé *assembleur*. Les lecteurs intéressés par ces problèmes sont très vivement engagés à consulter les ouvrages spécialisés, où ils les trouveront développés.

En effet, il existe de très nombreuses façons de procéder, et de très nombreux programmes dits « d'aide au développement », évoqués dans la figure 12 qui montre une façon de procéder générale.

Enfin, la figure 13 résume le mode de réalisation d'un logiciel d'application selon deux méthodes, l'une avec système de développement et l'autre, en temps partagé.



En deux circuits, le CODEC de Siemens (bicanal).

TÉLÉCOMMUNICATIONS : LES CODEC

Le développement espéré des télécommunications numériques a inspiré la création de nouveaux circuits spécialisés, codeurs-décodeurs, appelé CODEC par la contraction de ces deux mots. Ils réalisent les conversions analogique-numérique à l'émission, et inverse à la réception, afin de permettre des transmissions vocales en digital.

La plupart d'entre eux sont réalisés selon les technologies MOS bien qu'on en trouve également en I2L) ; c'est pourquoi ils vont être présentés ici, trouvant très naturellement leur place après les microprocesseurs. C'est qu'en effet, on recherche aussi bien une très grande densité d'intégration qu'une très faible consommation.

1. PRINCIPES

La modulation par codage d'impulsions, ou MIC (modulation par impulsions codées, ce qui se dit encore PCM en anglais, pour « pulse code modulation »), bien qu'ancienne, n'a pas encore connu toute l'exploitation qu'elle mérite.

Cette technique consiste à échantillonner un signal analogique et à convertir ensuite chaque échantillon en un code numérique, représenté par des impulsions. Ainsi, en téléphonie, une bande passante étroite de 3,4 kHz est échantillonnée au rythme de 8 kHz (car il faut plusieurs échantillons par période), et chaque échantillon est codé en numérique pour former un mot de 8 bits, à l'aide d'un convertisseur analogique-numérique (fig. 1). Ces informations série sont ensuite transmises par des voies habituelles.

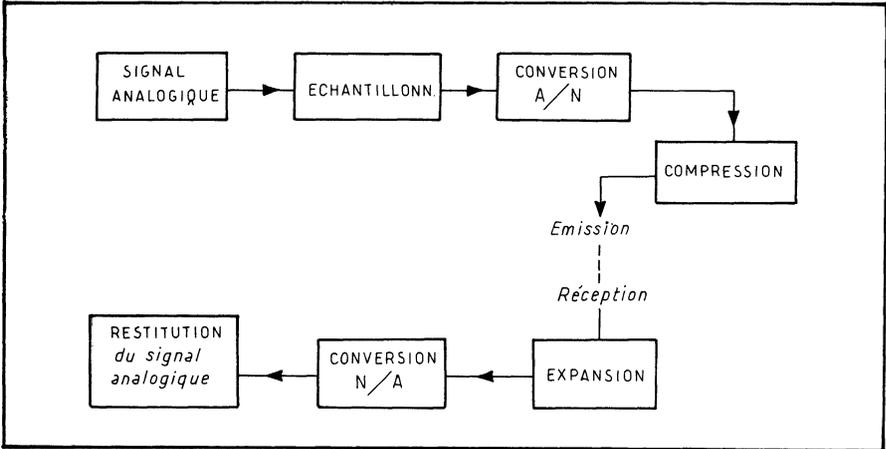


Fig. 1. — Principe d'intervention des CODEC (codeurs-décodeurs).

La compression-expansion

A la réception, elles subissent une transformation inverse afin de restituer l'information analogique dont elles sont porteuses. En outre, la transmission n'est assurée qu'après compression de la dynamique (avec expansion symétrique à la réception), et ce, selon des lois normalisées, approximativement logarithmiques. L'une est la loi dite μ 255 de la *Bell Telephone*, qui porte sur une compression de 72 dB; l'autre, adoptée en Europe et donc en France, est dite loi A et répond à la norme A 87-6 du CCITT; elle porte sur une compression de 66 dB (fig. 2 et 3).

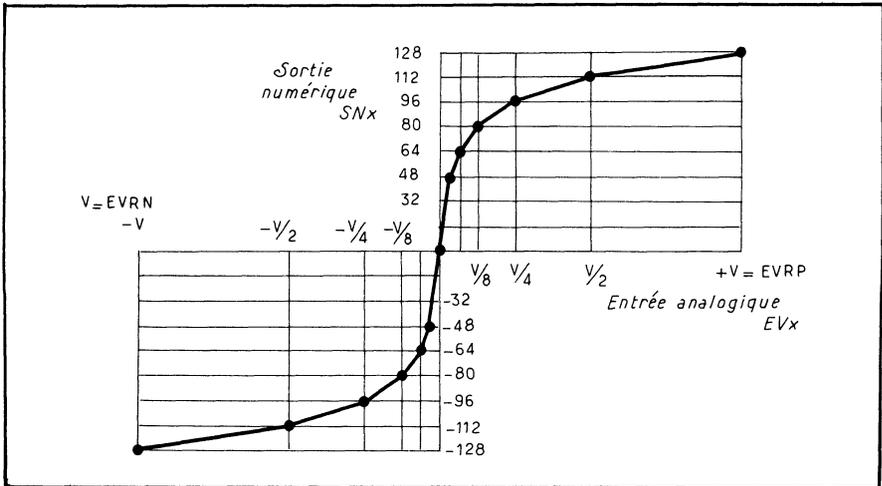


Fig. 2. — Courbe de transfert de la loi A de compression.

Pourquoi passe-t-on par une compression? Parce que l'on ne peut prélever les échantillons de l'onde analogique que sur 256 niveaux, si l'on utilise un codage sur 8 bits; en effet, $2^8 = 256$. De ce fait, la courbe d'origine ne pourra être qu'imparfaitement restituée, après conversion inverse à la réception (on veut reconstituer une fonction linéaire à l'aide de paliers sur 256 niveaux); l'erreur introduite est appelée *bruit de quantification*, et elle est fonction de l'intervalle entre deux niveaux successifs uniquement (et non de l'amplitude du signal). De ce fait, le rapport (*signal*) sur (*bruit de quantification*) est plus faible aux faibles niveaux d'amplitude (lorsqu'il n'y a pas compression).

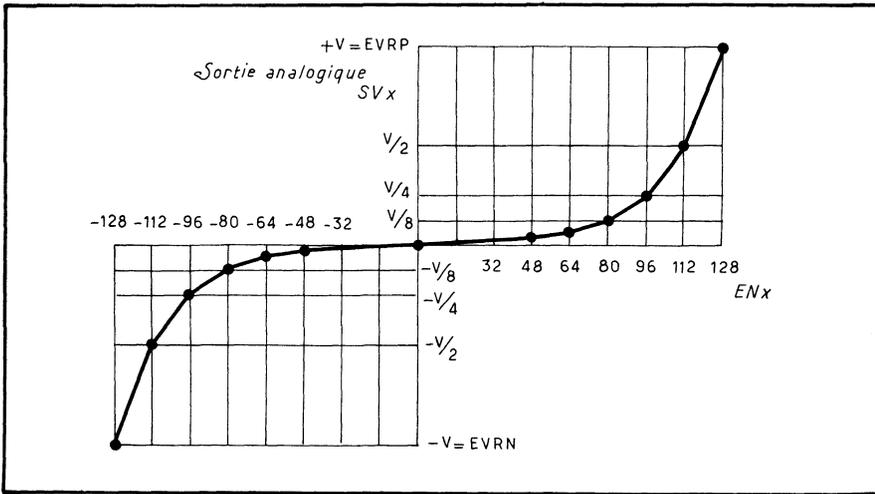


Fig. 3. — Loi A d'expansion, à la réception.

Afin d'obtenir un tel rapport *constant*, quelle que soit l'amplitude analogique, il faut que les faibles signaux rencontrent davantage de niveaux que les forts, ce qui fait immédiatement songer à une courbe logarithmique : c'est celle de compression du signal analogique avant codage. Ainsi, le rapport du signal sur l'erreur, ou le bruit de quantification, reste constant. Diverses recommandations relatives aux lois de compression-expansion ont été formulées par le CCITT; on pourra s'y référer (notamment, à la recommandation A 711 sur les *compandeurs*, mot qui résulte de la contraction de compresseur-expandeur, dans le Livre vert, vol. III, Lignes de transmission, 1973).

D'autre part, le codage par modulation d'impulsions est relatif soit aux *niveaux absolus* des échantillons du signal, c'est le MIC classique, soit à l'évolution du signal analogique et non plus à ses niveaux; il porte alors le nom de *modulation delta*.

Or, avant l'avènement de ces fonctions en circuits intégrés, le codage et le décodage se révélaient très complexes et c'est ce qui en a freiné le développement. Avec les progrès de l'intégration, on a vu naître de nouveaux circuits qui regroupent tout ou partie de ces fonctions afin de constituer ce qu'on appelle des CODEC, pour *CO*deurs-*DE*Codeurs.

2. COMMENT FONCTIONNE UN CODEC

De par ses fonctions, un CODEC doit donc recevoir le signal analogique, l'échantillonner, mettre en mémoire l'échantillon, assurer sa conversion en numérique, appliquer la loi de compression retenue, et l'émettre. En sens inverse, c'est-à-dire à la réception, il accueille un signal numérique, assure son expansion, puis sa conversion en analogique, et restitue les échantillons qui serviront à reconstituer le signal analogique d'origine. Cela, pour un fonctionnement en MIC classique, par exemple.

La figure 4 illustre cette démarche; elle représente le synoptique du CODEC, en MIC, de EFCIS, référencé CDL 400 et réalisé en MOS complémentaires sur substrat isolant saphir (CMOS/SOS).

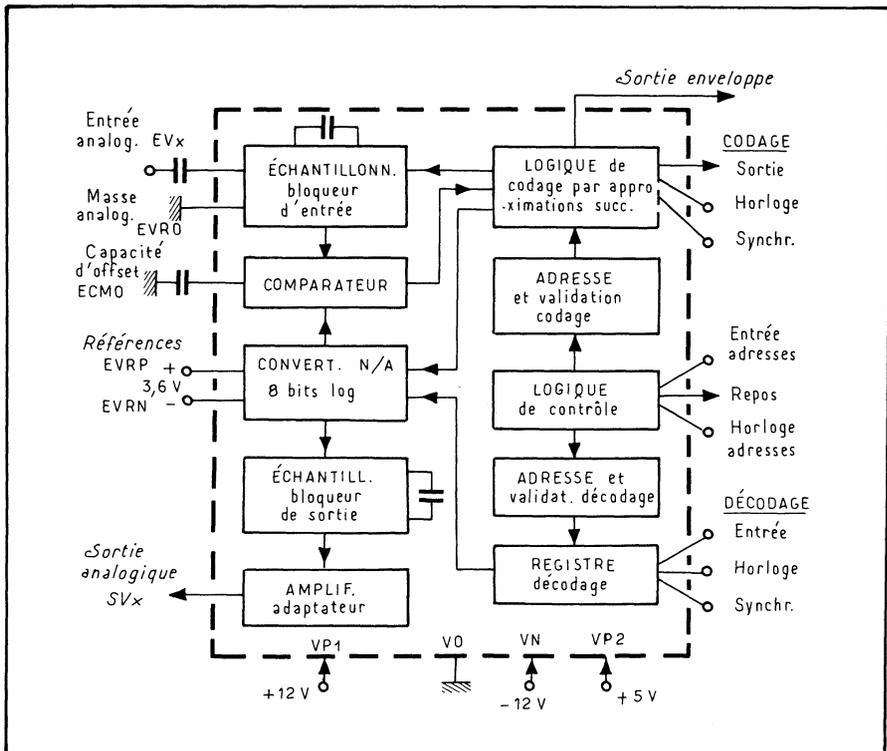


Fig. 4. — Synoptique du CODEC de EFCIS.

Le circuit effectue le codage du signal analogique d'entrée « EVx » sur 8 bits, suivant la caractéristique de conversion à compression de dynamique de la norme A du CCITT. Le résultat est disponible sur la sortie numérique série, 3 états 8 bits, SNx, dont le débit est défini par l'horloge de codage « EHC ». Le processus de conversion est initialisé par l'entrée synchro codage ESTC qui définit donc la fréquence d'échantillonnage (8 kHz à 24 kHz).

Parallèlement à l'opération de codage, le circuit effectue le décodage du signal numérique 8 bits présenté sur ENx, suivant la caractéristique de conversion à expansion de dynamique de la norme A 87,6, du CCITT. Le résultat est disponible

sur la sortie analogique SVx avec une résolution de 8 bits et une dynamique de 12 bits. Le processus de décodage est défini au sein de la séquence de codage; seule l'introduction du mot 8 bits sur ENx est totalement asynchrone, grâce à une horloge de décodage EHD et à une entrée synchro ESTD indépendantes du codage.

Le CODEC CDL 400 possède, en outre, deux registres 8 bits permettant de mémoriser l'introduction, par l'entrée d'adressage EADR et l'horloge d'adressage EHAD, du mode de fonctionnement choisi (repos, codage seul, décodage seul, codage et décodage), ainsi que des numéros de voies temporelles choisies pour l'émission et pour la réception (ces numéros peuvent être différents).

On voit qu'avec ce circuit, c'est le convertisseur numérique-analogique (qui intervient dans la conversion inverse selon le principe des approximations successives) qui assure la compression-expansion logarithmique.

Un autre exemple de synoptique est donné figure 5, pour les AY-3-9900 (loi A de compression) et 9910 (loi μ de *General Instrument*). Ici, des filtres ont été prévus sur les voies analogiques afin de limiter la bande passante, et le codage est du type delta. Le débit (vers le bus numérique) est de 2 048 kbits par seconde,

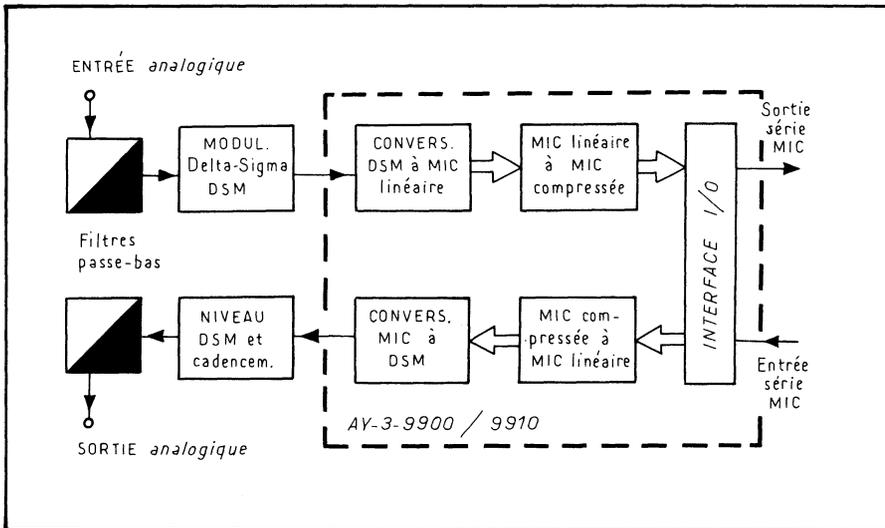


Fig. 5. — Synoptique du CODEC de G.I. et introduction dans la voie de télécommunications.

valeur très souvent adoptée car elle permet le multiplexage de 32 voies de 64 kbits par seconde, ce qui correspond enfin à un échantillonnage à 8 kHz avec un mot de un octet (8 bits) par échantillon. La transmission numérique se fait en série, de 64 kbits à 2 048 kbits par seconde (valeurs normalisées).

En fait, ici, l'échantillonnage est assuré à 2 048 kHz car en modulation delta, il suffit de coder l'évolution de l'onde analogique, le signe de sa dérivée, et par conséquent un unique bit est suffisant. La conversion analogique à numérique est donc bien plus simple. La compression ramènera ensuite cette dynamique à la valeur normalisée de 64 kbits/seconde. Cet AY-3-9900 n'assure donc que la conversion modulation delta en MIC, et vice-versa; il lui faudra ainsi des circuits d'accompagnement.

3. L'EXEMPLE DE CODEC TYPES

Les CODEC commercialisés sous ce nom revêtiront, on s'en apercevra, les structures les plus diverses, des plus simples aux plus complexes; parmi eux, l'un des plus intéressants est probablement le S 2563 annoncé par AMI en 1978. Conçu en CMOS, il comporte ses filtres de bande, mais se présente en deux circuits, l'un pour l'émission et l'autre pour la réception; la voie émission est décrite fonctionnellement dans la figure 6 : deux filtres jouent le rôle, l'un de passe-haut (Tchebyshev, d'ordre 3) et l'autre de passe-bas; la conversion A/D fait appel à la technique à charges capacitives; une boucle à phase asservie (« PLL » : phase locked loop) crée les fréquences de référence.

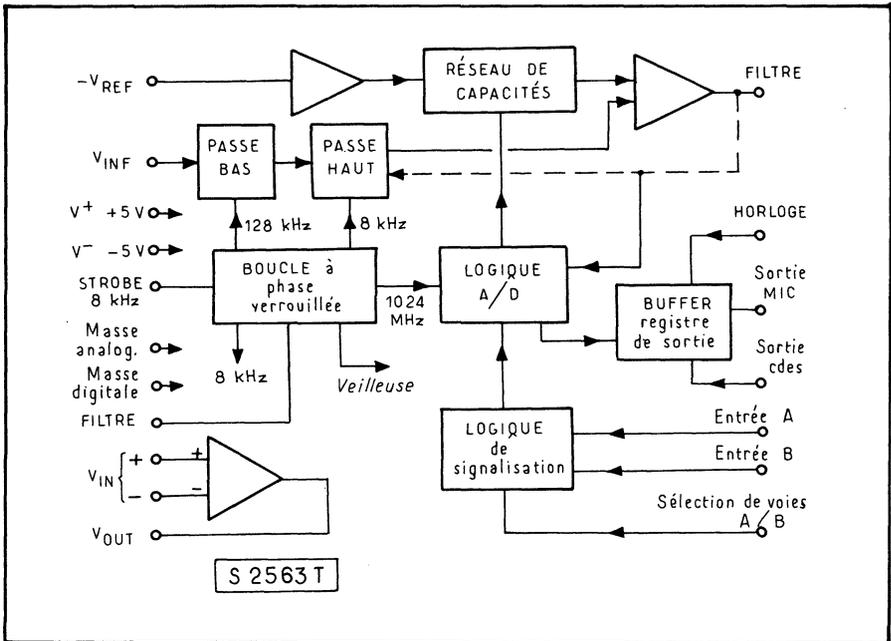


Fig. 6. — Le S 2563T de AMI, annoncé en 1978, comprendrait ses filtres et disposerait de deux voies d'entrée.

C'est également à une conversion par échelle de capacités que recourt *Siliconix*, avec ses tout récents DF 331/332. Là encore, il s'agit de CODEC « mono-voies » (un seul canal), en deux circuits, l'un pour l'émission et l'autre pour la réception et en boîtiers 16 broches.

RTC-Signetics préfère la technologie I²L et assure que les performances obtenues sont supérieures. Son ST 100 (fig. 7) exige cependant trois tensions d'alimentation, mais fournit des informations supplémentaires qu'ignorent généralement les autres modèles (information bicanale selon les spécifications D3 de la *Bell*) et supprime le code « zéro » en introduisant des 1 supplémentaires après des chaînes de zéros, afin de ne pas perdre la synchronisation.

Bivoies et avec une formule technologique mixte, les SM61A (en NMOS) et S291 (bipolaire) de *Siemens* assurent les mêmes fonctions (fig. 8). Ils appliquent

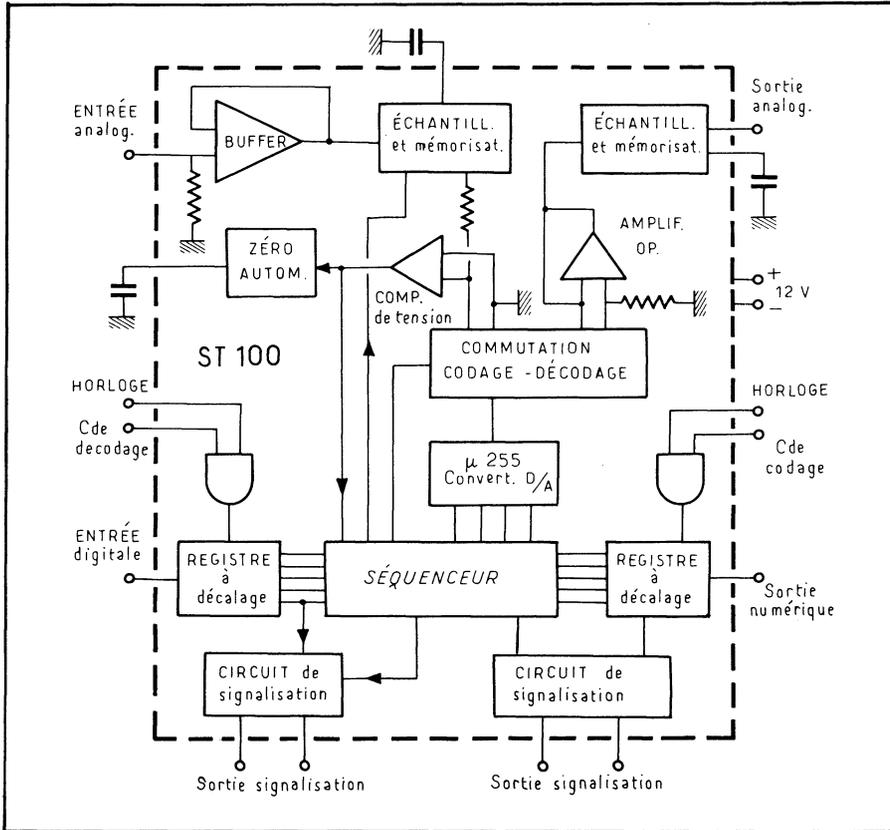


Fig. 7. — Synoptique du ST 100 de RTC-Signetics, en technologie I²L.

la loi A mais ne comportent pas les filtres, et fonctionnent en duplex intégral avec deux circuits téléphoniques indépendants sur 4 fils.

Par contre, *Intel* qui a misé sur le NMOS, propose un CODEC et un seul circuit assurant et le codage, et le décodage. Ce qui est également le cas de *Mostek*, de *Signetics* déjà cité, de *Motorola* et de *GI* (voir notre tableau).

Ajoutons que c'est *Intel* qui a lancé la première version monolithique d'un Codec monovoie au début de 1978, avec les circuits 2911 et 2910. Ces circuits incorporent une logique de calcul automatique des intervalles de temps pour transmission et réception ainsi qu'une tension de référence pour les processus de codage-décodage. Fin 1978, cette même société annonçait les filtres transmission-réception sous forme monolithique également, le circuit 2912.

Monovoies et multivoies

On remarque que les CODEC proposés relèvent de très nombreuses variantes techniques, rien n'étant réellement figé à ce jour. L'un des problèmes débattus réside dans le choix entre monovoie et multivoie. Dans le premier cas, on utilise

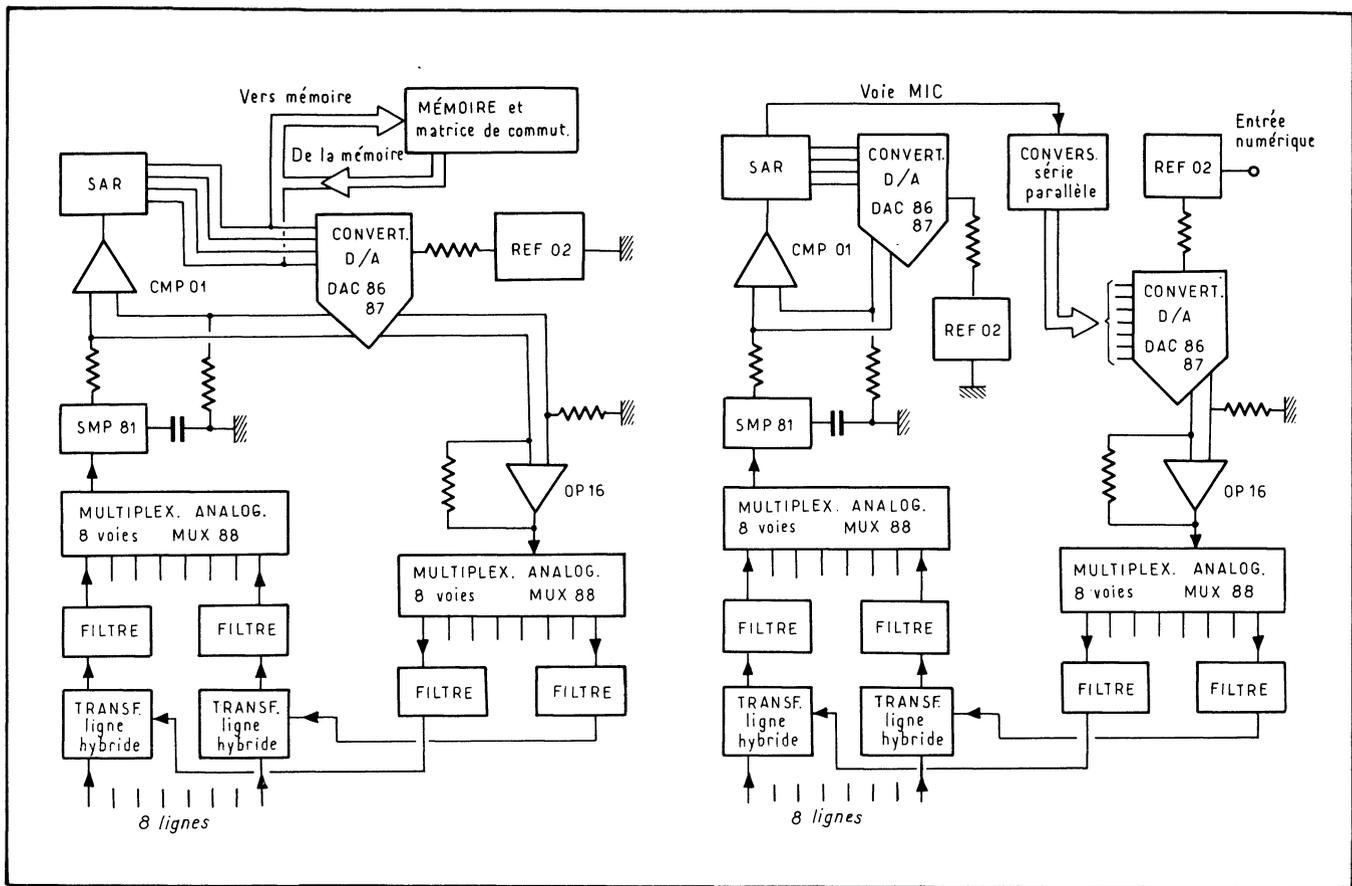


Fig. 9. — Les multivoies de Precision Monolithic (PMI), du groupe Ohmic-Bourns. C'est la solution actuellement la plus économique lorsqu'il y a plusieurs voies à multiplexer. En (a), deux convertisseurs D/A du type COMDAC interviennent; il n'en subsiste plus que un en (b).

Quelques CODEC

Société	Référence	Nombre de circuits	Monoivoie Multivoie	Compression-exp.		Filtres inclus
				Loi A	Loi μ	
AMD		2	Mono	•	•	Non
AMI	S2900	1	Mono	•		Oui
	S2901	1	Mono	•		Oui
	S2902	1	Mono		•	Oui
	S2903	1	Mono		•	Oui
	S2563T	1	Bi	•	•	Oui
EFCIS	CDL400	1	Mono	•		Non
General Instruments (GI)	AY3-9900	1	Mono	•	•	Non
Intel	2910	1	Mono		•	Non
	2911	1	Mono	•		Non
Mostek	MK5150	1	Mono		•	Non
	MK5116	1	Mono		•	Non
	MK5155	1	Mono	•		Non
Motorola	MC14407	1	Mono	•	•	Non
	MC14406	1	Mono		•	Non
National Semi-conductor	TP3001	1	Mono		•	Non
	TP3002	1	Mono	•		Non
PMI	DAC87	—	Multi			Non
	DAC86	—	Multi	•	•	Non
Siemens	SM 61A S 291	} 2	Mono	•		Non
Signetics/ RTC	ST 100	1	Mono		•	Non
Siliconix	DF331	1	Mono		•	Non
	DF332	1	Mono		•	Non
	DF334	1	Mono		•	Non
	DF341	1	Mono	•		Non
	DF342	1	Mono	•		Non

présentés par les fabricants

Technologie	Brochage	Consommation (mW) activité/veille	Tensions d'alimentation	Observations
NMOS et Bipolaire	24			Loi A ou μ programmable par broche
CMOS	18	200/25	± 5 V	Codeur uniquement Décodeur uniquement Codeur uniquement Décodeur uniquement Annoncé en 1978
CMOS	16	200/25	± 5 V	
CMOS	18	200/25	± 5 V	
CMOS	16	200/25	± 5 V	
CMOS	18		± 5	
CMOS/SOS	22	40	+5 V, ± 12 V	
NMOS	24	300	+9 V, +5 V	Loi μ ou A sélectionnée par une broche
NMOS	24	220/110	+12 V, ± 5 V	
NMOS	22	220/110		
CMOS	24	30	± 5 V	
CMOS	16	30	± 5 V	
CMOS	16	30	± 5 V	
CMOS	28	80/0,5	10, -16 V	Loi A ou μ sélectionnée par broche
CMOS	24	80/0,5	10, -16 V	
CMOS/BIFET CMOS/BIFET				
Bipolaire	16	140	± 10 à 15 V ± 10 à 15 V	} Convertisseurs D/A rapide uniquement, associés à d'autres circuits
Bipolaire	16	140		
NMOS	28			
Bipolaire	24			
I2L	24	375/50	± 12 , +5 V	
CMOS	14	45	$\pm 7,5$ V	Codeur Décodeur Décodeur Codeur Décodeur
CMOS	14	45	$\pm 7,5$ V	
CMOS	14	45	$\pm 7,5$ V	
CMOS	14	45	$\pm 7,5$ V	
CMOS	14	45	$\pm 7,5$ V	

un CODEC par voie, dans le second, les informations analogiques passent auparavant par un multiplexeur avant d'attaquer un unique CODEC (donc un pour toutes les voies). Si la première solution est plus souple, la seconde est plus économique actuellement, mais exige un convertisseur analogique-numérique, qui est l'âme du système, très rapide, ce qui n'est pas à l'avantage des technologies MOS courantes. La discussion porte sur bien d'autres avantages et inconvénients de ces deux formules.

Le champion du multivoies, qui est également un pionnier en matière de systèmes CODEC, est *Precision Monolithic INC*, ou *PMI*, du groupe *Ohmic-Bourns*. Les ensembles qu'elle propose sont d'autant plus intéressants qu'elle en a chiffré le prix; ils sont bâtis autour de convertisseurs bipolaires très rapides, les « *COMDAC* », qui assurent tout à la fois la compression et la conversion, la figure 9 donnant le synoptique des deux structures de systèmes proposés; dans l'une, les voies émission et réception recourent au même *COMDAC* alors qu'il en faut deux dans l'autre, pour 8 lignes multiplexées. Les quelque 11 boîtiers intégrés nécessaires à la formule la plus économique ramène l'investissement à \$ 1,67 par ligne, soit à peu près 8 F. Selon l'autre formule, il en coûterait \$ 2,36 par ligne, soit environ 11 F.

Consommation. — Un autre point très sensible réside dans la consommation, le courant nécessaire pouvant n'être fourni que par les lignes téléphoniques, par exemple. De ce fait, l'avantage reviendrait aux technologies peu gourmandes en énergie, *I²L* et *CMOS*. Le *Codec* de *Motorola* (*MC 14406/14407*), en *CMOS*, ne dissipe ainsi que 80 mW en activité et 0,5 mW à l'état de veille, contre plusieurs centaines de milliwatts pour les autres.

4. MARCHÉS ET APPLICATIONS

S'il est vrai que les applications essentielles des *CODEC* relèvent de la téléphonie, il est non moins exact qu'ils peuvent s'appliquer efficacement à toutes transmissions numériques d'informations industrielles, tout comme à la commande de servomécanisme, avec le bénéfice de l'excellente précision relative, constante, que procure la compression logarithmique.

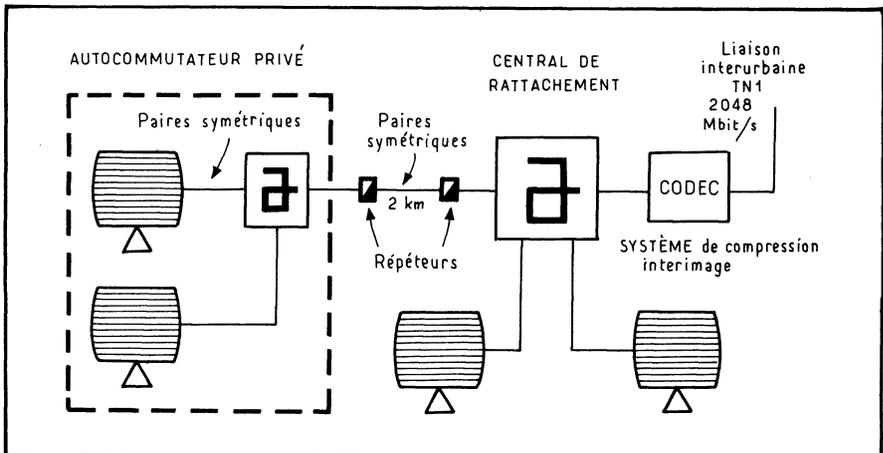


Fig. 10. — Dès 1975, le CNET proposait un visiophone (service de transmission d'images animées en temps réel exploité sur réseau téléphonique) dans lequel la voie son recourt à un *CODEC*, selon ce synoptique.

En téléphonie, toutes les grandes sociétés françaises ont procédé à l'évaluation des CODEC. On attend surtout les décisions de principe des Administrations. En fait, plusieurs applications des CODEC sont déjà en service; dès 1975, le visiophone du *CNET* utilisait un CODEC sur sa voie son (fig. 10). Des applications industrielles sont étudiées pour les transmissions vocales dans des milieux très perturbés, mines par exemple. On pense également qu'en téléphonie classique, ce seront les réseaux à centraux privés qui les exploiteront tout d'abord.

Le marché devrait donc être immense. Peu significatif en 1978, selon *Motrola*, il devrait se monter d'après ses estimations à 3,5 millions de CODEC en 1979, et à 7 millions en 1983, alors que *Mostek* en prévoit de 3 à 5 millions pour 1980.

L'une des premières sociétés à introduire des CODEC, *PMI* déjà citée, annonçait avoir vendu plus de 500 000 CODEC en deux ans, dont 2 000 en France. Cette indication est d'autant plus précieuse que l'on connaît la répugnance des fabricants à faire état de leurs ventes.

CHAPITRE XVIII

APPLICATIONS LINÉAIRES

Les 4/5 des circuits intégrés s'appliquent au numérique. Il en reste donc 20 % environ dont la destination est linéaire. Dans ce domaine également, les MOS ont commencé une offensive qui s'est traduite, entre autres, par la réalisation d'amplificateurs opérationnels « mixtes » parmi les plus performants qui soient.

1. LE MOS EN LINÉAIRE

Si l'on reprend les caractéristiques typiques du MOS, on voit que dans la zone B (saturée), où le transistor se comporte comme une source de courant constant, une variation de la tension de commande de gate V_{GS} se traduira par une variation du courant drain I_D . La pente est alors dI_D/dV_{GS} , ou son inverse, la transconductance, est :

$$G_m = \frac{dI_D}{dV_{GS}}$$

Un tel MOS peut donc servir d'amplificateur.

2. APPLICATIONS AF

Amplificateur à basses fréquences

La polarisation du MOS, en linéaire, peut être assurée par une simple résistance R_P montée en contre-réaction entre drain et gate (fig. 1 a). En l'absence de courant I_P à travers R_P s'établit l'état $V_{GS} = V_{DS}$ pour les tensions continues. La courbe correspondante sur le réseau habituel I_D fonction de V_{DS} (en b) coupe la droite de charge due à R_L en un point qui est le point de fonctionnement en courant continu.

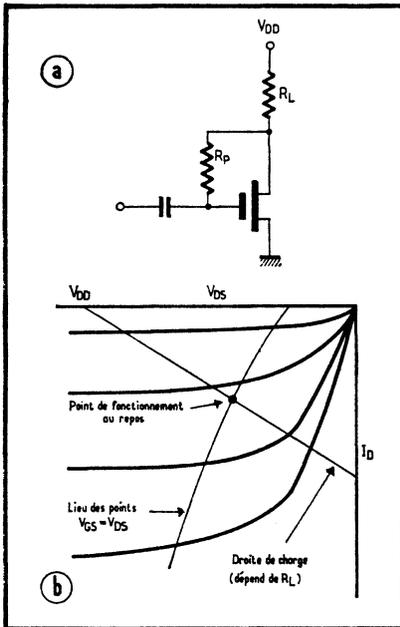


Fig. 1. — Polarisation d'un MOS utilisé en amplificateur linéaire (a). Courbes caractéristiques et point de fonctionnement au repos (b). Principe du couplage direct des MOS (c).

On constatera que la valeur de R_P ne joue pas sur la position du point de fonctionnement, lequel ne dépend que de la valeur de la charge. En effet, aussi grande que soit R_P , elle reste faible devant l'impédance d'entrée du MOS (à moins de recourir aux centaines de mégohms...).

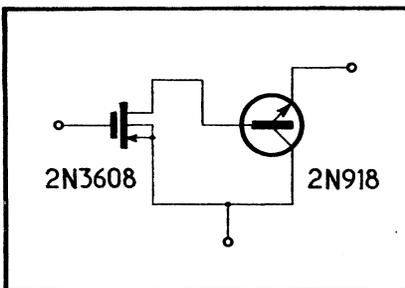


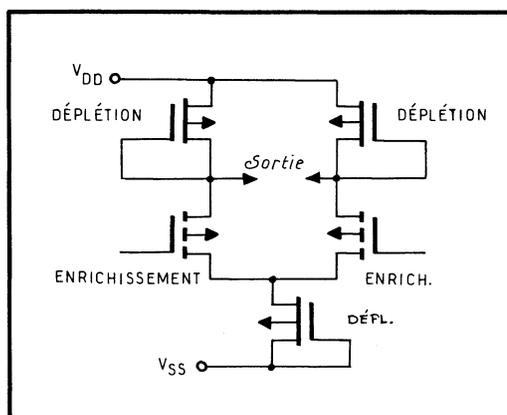
Fig. 2. — Un transistor bipolaire associé à un MOS accroît la pente résultante.

La charge est soit une résistance, soit un autre MOS; dans ce dernier cas, le MOS en charge travaillera de préférence en régime saturé pour que la linéarité reste bonne sur la plus large gamme de tensions : en effet, les non-linéarités des deux MOS en série tendent alors à s'annuler. La linéarité globale est alors supérieure à celle obtenue avec une résistance passive en charge.

MOS avec émetteur suiveur

Pour accroître la transconductance du MOS, on lui adjoint parfois un transistor bipolaire monté en émetteur-suiveur, comme le montre la figure 2. En consé-

Fig. 3. — Amplificateur différentiel à charges à déplétion.



quence, la pente du MOS est multipliée par le gain en courant H_{FE} du bipolaire : si la pente est de 1 000 micromhos et le gain de 100, la pente résultante sera de 0,1 mho. Physiquement, on bénéficie de l'identité des concentrations relatives des diffusions P du MOS et de base du bipolaire.

Un CI de ce type a été réalisé par *RTC-La Radiotechnique-Compelec* sous l'appellation TAA 320.

3. STRUCTURE DIFFÉRENTIELLE

Dans la figure 3, on a représenté un étage différentiel à MOS monocanaux, mais utilisant des charges à déplétion (tel qu'utilisé en amplificateur de lecture dans le circuit 4007 *Mostek*). Un tel montage, à plusieurs étages, offre une sensibilité meilleure que 100 mV, et se prête à la réalisation d'oscillateurs sinusoïdaux, d'amplificateurs, de détecteurs de zéro, etc.

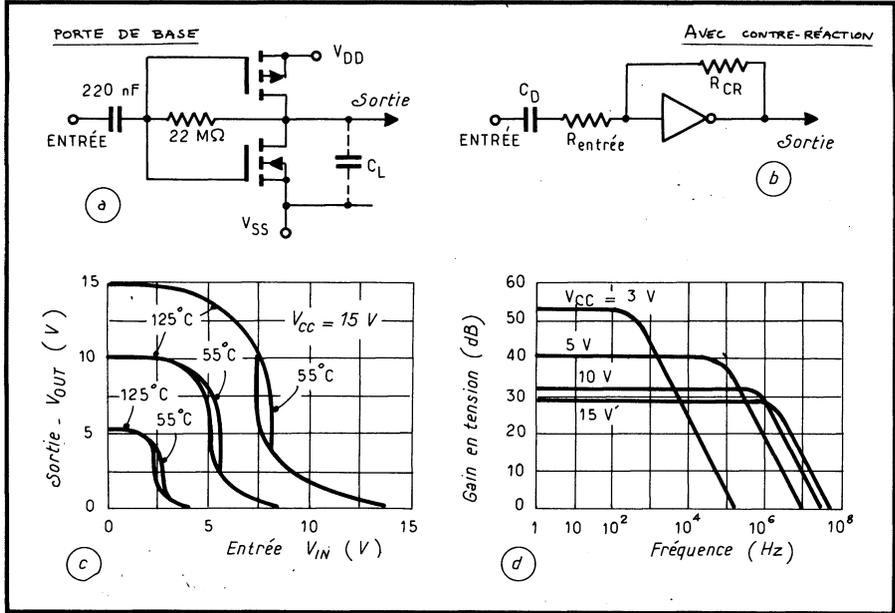


Fig. 4. — L'amplificateur à CMOS.

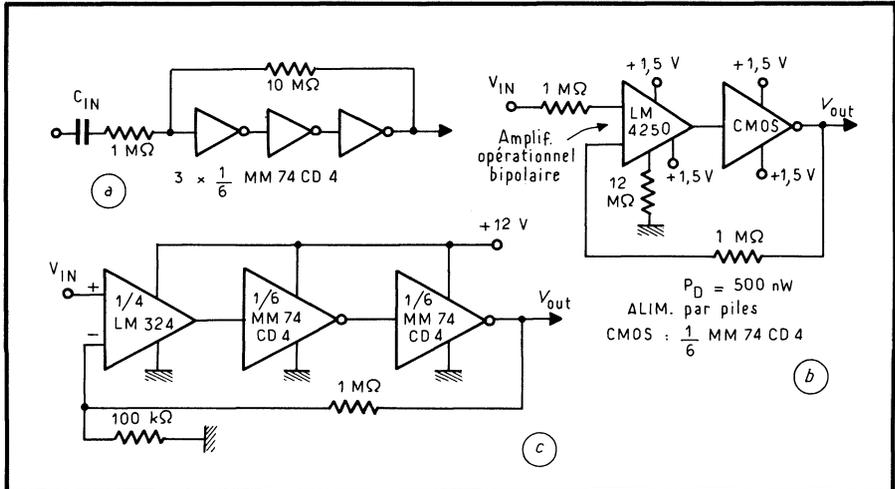


Fig. 5. — Trois inverseurs en cascade constituent un amplificateur linéaire, à gain de 10 ici en raison de la valeur des résistances (a). Amplificateur CMOS attaqué par un bipolaire (b), et montage d'inverseurs avec unique tension d'alimentation (c).

4. MONTAGES A CMOS

Le principe de la polarisation d'un étage à CMOS est donné figure 4 a, ce qui se traduit par le schéma b pour un inverseur type 74 C 04 de NS, donné avec ses courbes de transfert (c). En d apparaissent les courbes typiques de gain en tension obtenues en fonction de la fréquence.

Un amplificateur de ce type à 3 étages en cascade est représenté figure 5 a. On peut utiliser un inverseur en post-amplificateur pour amplificateur opérationnel (en b); l'avantage réside dans l'excursion complète de la sortie entre les tensions d'alimentation. Dans le schéma c, les étages sont alimentés à partir d'une unique source de tension; avec l'amplificateur opérationnel bipolaire LM 324, le gain global est de 160 dB.

5. LA BOUCLE A PHASE ASSERVIE

La boucle à phase asservie, ou PLL (de *phase locked loop*) a connu un nouveau développement avec les circuits intégrés. La fonction d'une telle boucle consiste à comparer la phase d'une onde incidente à celle d'une référence interne dont la fréquence est asservie à l'incidente. Elle s'applique de ce fait à la démodulation, au décodage, mais aussi à la multiplication et division de fréquence, etc., car son mode de traitement est numérique.

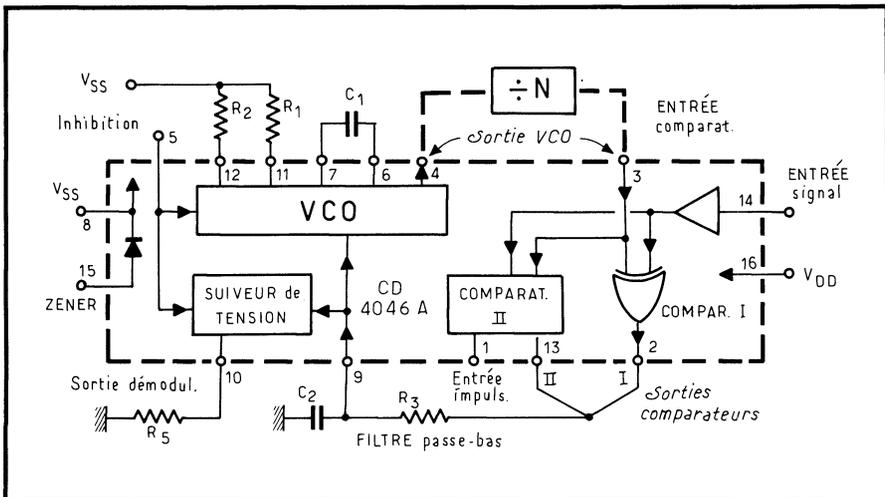


Fig. 6. — Synoptique du CD 4046, boucle à verrouillage de phase réalisée par RCA en MOS complémentaires. Ce CI consomme moins de 1 mW.

De telles PLL ont surtout été réalisées en bipolaire. Avec les CI à MOS complémentaires et dans l'état actuel de la technique, la fréquence de travail est encore limitée : à 500 kHz avec la boucle à phase asservie CD 4046 A de RCA.

Par contre, l'économie en consommation est évidente : de 100 mW environ, elle passe à quelque 600 μ W (sous 6 V et à 10 kHz).

Le synoptique de ce CI est donné figure 6. On trouve un VCO, ou oscillateur commandé par une tension, associé à un double comparateur de phase, du type numérique. Une Zener stabilise l'alimentation. Le VCO est accordé par un condensateur externe C_1 et peut être relié aux comparateurs soit directement, soit par l'intermédiaire d'un diviseur de fréquence extérieur lui-même. Enfin, l'habituel filtre passe-bas réapplique au VCO la tension de correction.

6. LES AMPLIFICATEURS OPÉRATIONNELS

C'est peut-être là où on les attendait le moins que les MOS ont remporté leurs plus brillants succès : dans les amplificateurs opérationnels. Dominé longtemps sans partage par les bipolaires et avec pour première étoile le célèbre 709, le marché des « amplis op » disposait d'une panoplie d'éléments dont les performances semblaient atteindre des limites difficiles à dépasser. Or, en 1974 et 1975 sont apparus :

- les amplificateurs à chopper de *Harris Semiconductor*, puis *Texas*;
- les amplificateurs à entrée MOS, étages intermédiaires bipolaires et sortie CMOS de *RCA*;
- les amplificateurs à JFET.

L'amplificateur à chopper HA 2900

L'amplificateur opérationnel à chopper d'entrée HA 2900 de *Harris*, probablement l'élément monolithique le plus performant de ce type, combine des MOS et des bipolaires à l'isolement diélectrique.

Sa stabilité est exceptionnelle : l'offset est de 0,1 nA à 25 °C pour une dérive de 1 pA/°C, ou de 50 μ V pour 0,2 μ V/°C. Cela, parce que les étages d'amplification travaillent en alternatif, ce qui permet d'éliminer la dérive continue due à la température.

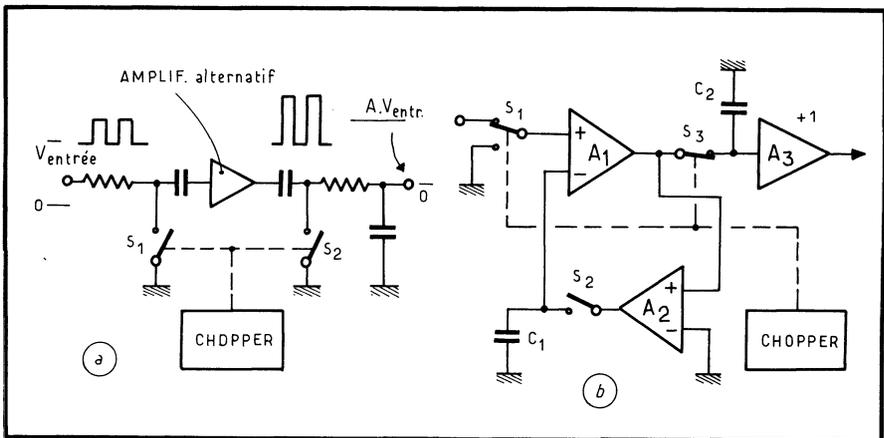


Fig. 7. — Chopper classique (a); celui du HA 2900 (b)

En effet, un « découpeur », ou chopper, découpe à l'entrée les signaux selon un principe très classique (fig. 7 a), mais modifié pour le HA 2900 comme le montre la figure b : l'amplificateur est périodiquement déconnecté de l'entrée et on profite de ces intervalles pour annuler l'offset.

Si S_1 et S_3 ferment les circuits comme sur la figure, la chaîne d'amplification A_1 , A_3 fonctionne normalement. Si S_1 , S_2 et S_3 inversent leur position (par rapport

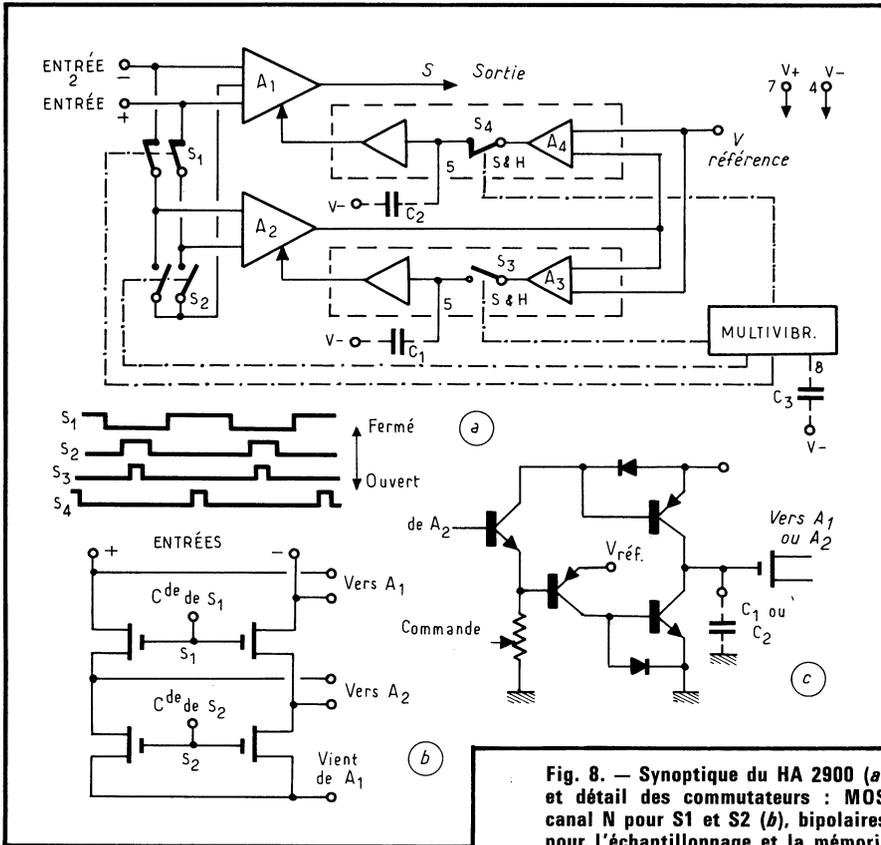


Fig. 8. — Synoptique du HA 2900 (a) et détail des commutateurs : MOS canal N pour S_1 et S_2 (b), bipolaires pour l'échantillonnage et la mémorisation (c).

à celle du schéma), l'entrée de A_1 est à la masse et A_2 impose une tension de sortie nulle à A_1 ; l'interrupteur S_2 et le condensateur C_1 constituent un circuit d'échantillonnage et mémorisation qui stocke la tension nécessaire à l'annulation du décalage; celle-ci reste appliquée à A_1 quand S_2 repasse en position haute.

Enfin, S_3 , C_2 et S_5 forment un second circuit d'échantillonnage et maintien qui mémorise la sortie de A_1 , pendant l'opération d'annulation du décalage. Ainsi, la sortie restituée par A_3 , de gain unitaire, est continue.

Le synoptique complet du CI est donné figure 8 a, et ses étages d'entrée, figure b et c : en (b) sont les N/MOS qui jouent le rôle de S_1 et S_2 , et en (c), les

commutateurs bipolaires des circuits d'échantillonnage et mémorisation. Le slew-rate du HA 2900 est de $2,5 \text{ V}/\mu\text{s}$ et la bande passante au gain unité est de 3 MHz typiquement.

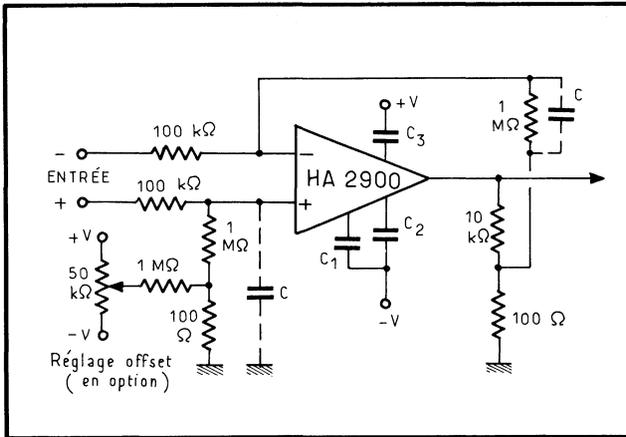


Fig. 9. — Amplificateur à gain élevé réalisé avec un HA 2900.

La figure 9 représente un montage typique différentiel du HA 2900 avec gain de 1000 et bande passante de 2 kHz. Cet élément s'applique bien évidemment aux quelques dizaines de schémas habituels aux amplificateurs opérationnels, mais en pratique, le montage ci-dessus est préconisé pour éliminer le bruit du chopper et réduire la dérive. Le bruit total ramené à l'entrée est ainsi de $30 \mu\text{V}$.

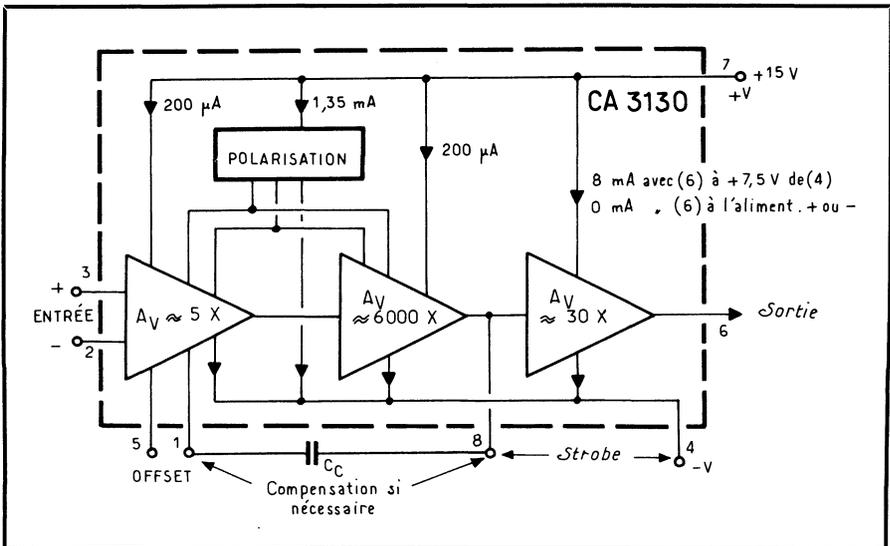


Fig. 10. — Synoptique du CA 3130.

L'amplificateur opérationnel BIMOS

Pour réquière encore le courant d'entrée, il faut recourir à des transistors MOS. C'est ce qui a amené RCA à proposer le premier amplificateur opérationnel à entrées sur P/MOS, étages intermédiaires bipolaires, et sorties sur C/MOS : le CA 3130.

Grâce aux PMOS, l'impédance d'entrée est très élevée et le courant d'entrée très faible, et grâce aux CMOS de sortie, l'excursion de tension avoisine, à quelque millivolts près, les tensions d'alimentation, celles-ci pouvant être simples (5 à 16 V) ou symétriques ($\pm 2,5$ à ± 8 V). La compensation des déphasages s'effectue à l'aide d'un unique condensateur externe, et l'offset à l'aide d'un potentiomètre de 100 Ω . La figure 10 donne le synoptique du CA 3130, présenté en boîtier métallique type TO-5 à 8 connexions.

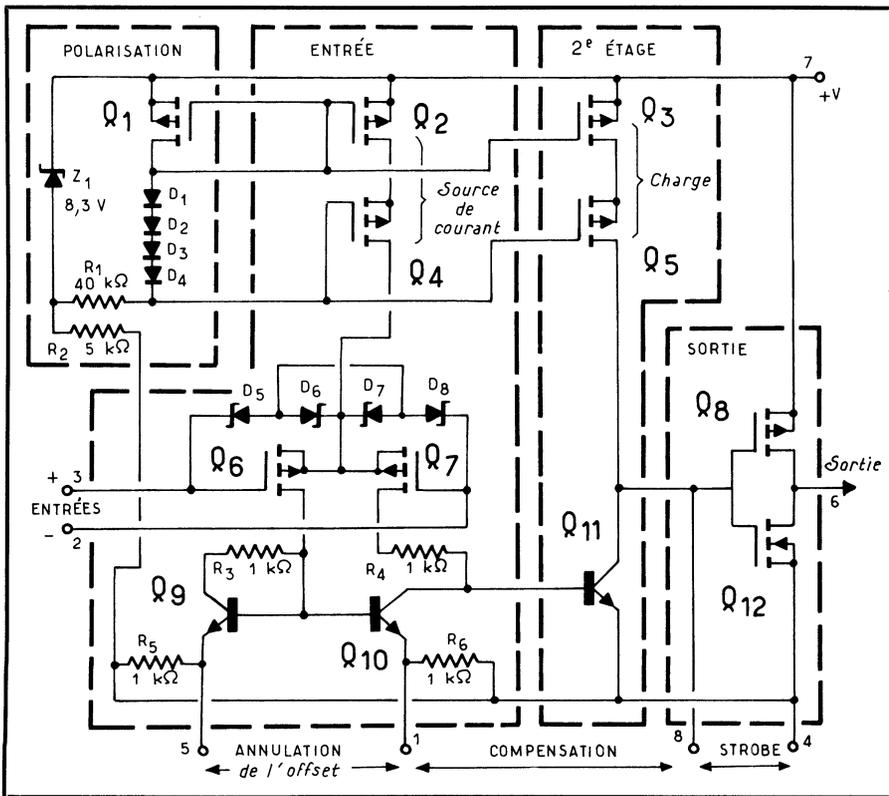


Fig. 11. — Schéma électrique du CA 3130.

La figure 11 représente son schéma électrique : à l'entrée, la tension peut être inférieure de 0,5 V à la tension négative d'alimentation. Le gain des PMOS est de l'ordre de 5, mais les bipolaires les suivent avec un gain A_v de 6 000.

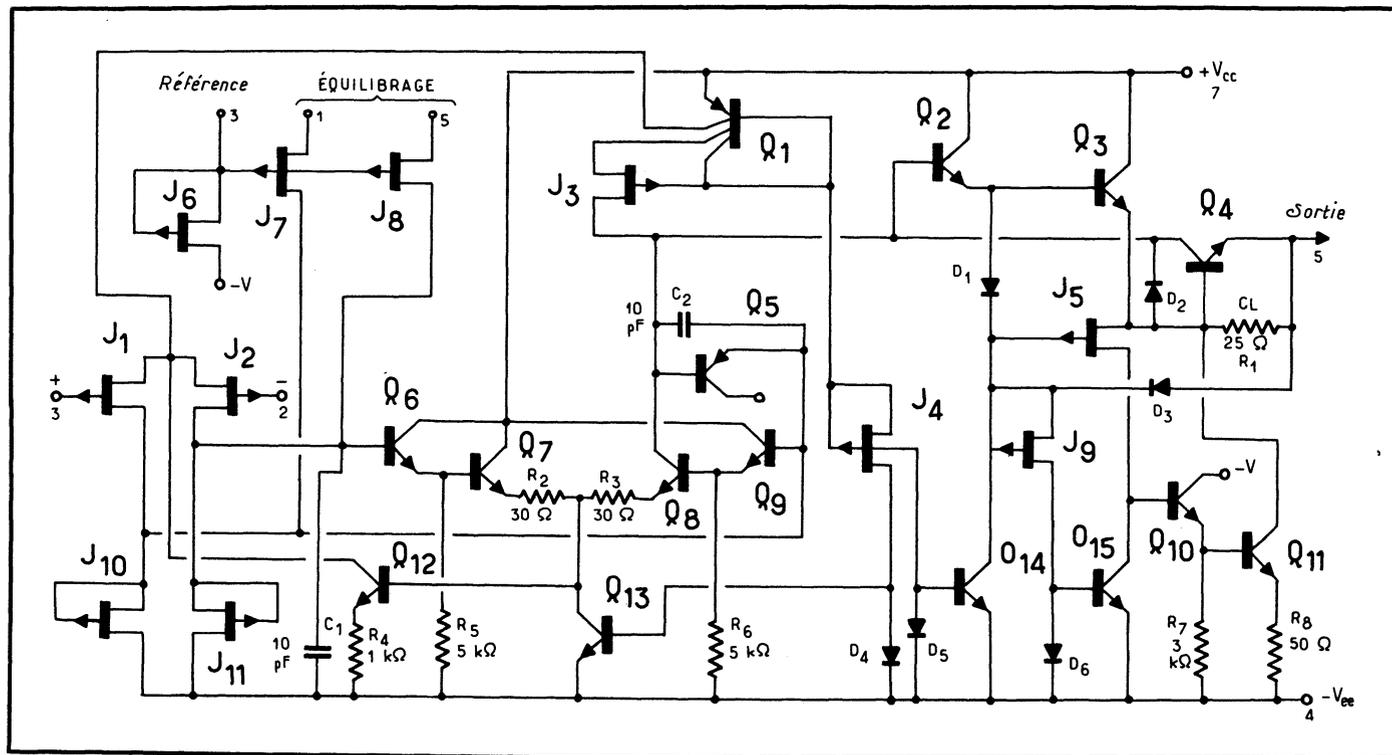


Fig. 12. — Schéma du « bifet » LF 156.

Fig. 13. — Générateur de signaux carrés à 0,5 Hz avec amplificateur opérationnel BIFET TL081.

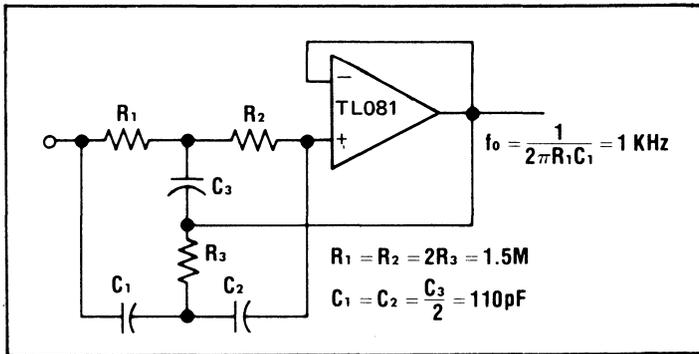
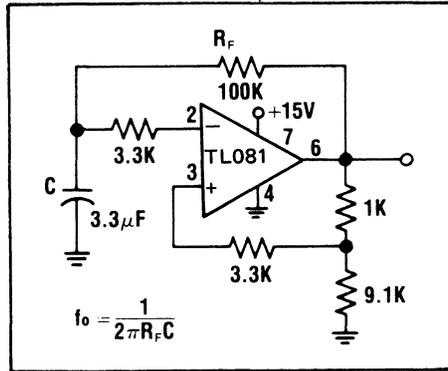


Fig. 14. — Circuit bouchon à Q élevé.

A propos du schéma électrique, on notera que l'entrée se fait sur une paire différentielle de PMOS qui attaque une paire bipolaire montée en « miroir » de courant pour justifier l'emploi d'une unique tension d'alimentation. De ce fait, l'entrée ne consomme que 5 pA et son impédance différentielle est de 1 000 MΩ.

Cette paire transpose le différentiel et attaque sur une voie le bipolaire Q₁₁, avec sa charge cascode Q₃ et Q₅ (des PMOS). Les PMOS Q₂ et Q₄ constituent, eux, une source de courant constant pour l'étage d'entrée, protégé par les diodes D₅ à D₈. La sortie est une porte CMOS classique, capable de piloter 22 mA; elle est auto-protégée contre les courts-circuits, ce qui n'est jamais le cas en bipolaire puisqu'il faut, là, adjoindre d'éventuels systèmes de protection complémentaires. En effet, le courant de sortie est auto-limité par la tension gate-source. De plus, le courant de court-circuit décroît lorsque la température croît.

Comme amplificateur opérationnel, ou comme comparateur, le CA 3130 n'exige qu'une seule tension d'alimentation mais est cependant capable de fournir une tension de sortie très proche de la tension négative. Cela n'est pas le cas avec les comparateurs bipolaires qui, dans les mêmes conditions, ne peuvent aller en dessous de 0,3 V. Il peut servir de détecteur de seuil précis, de convertisseur courant-tension, d'amplificateur classique, etc. Il préfigure les circuits mixtes analogique-logique de l'avenir où le traitement logique pourrait être assuré exclusivement par les C/MOS.

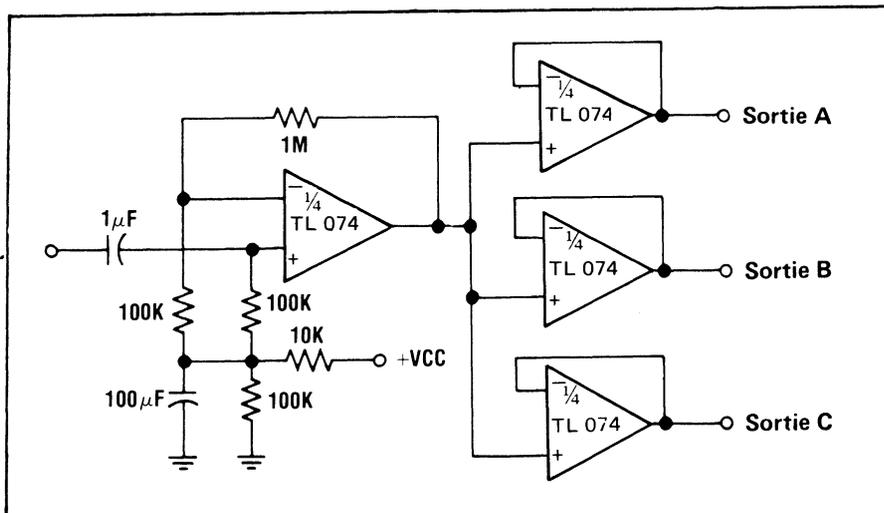


Fig. 15. — Montage amplificateur de distribution audio.

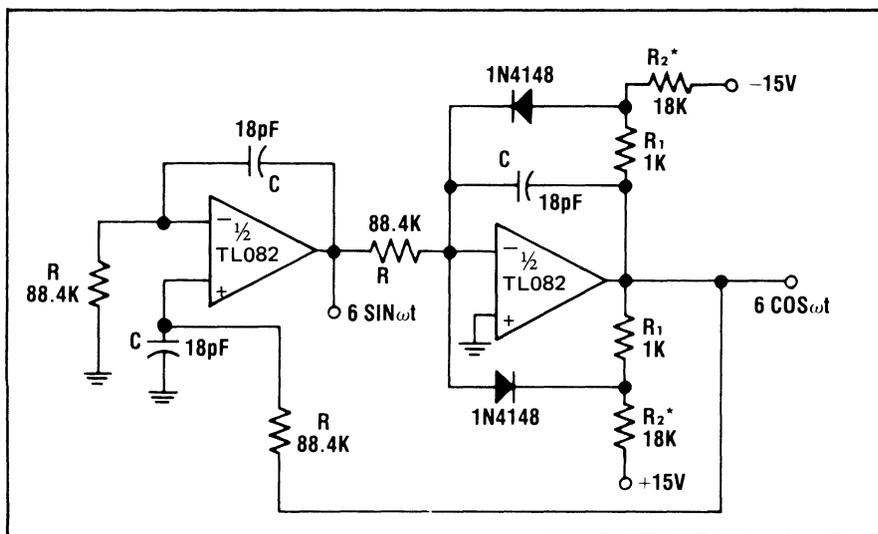


Fig. 16. — Oscillateur à quadrature avec $F_0 = 100\text{ kHz}$. Les résistances R_2 doivent être ajustées pour obtenir une sortie symétrique.

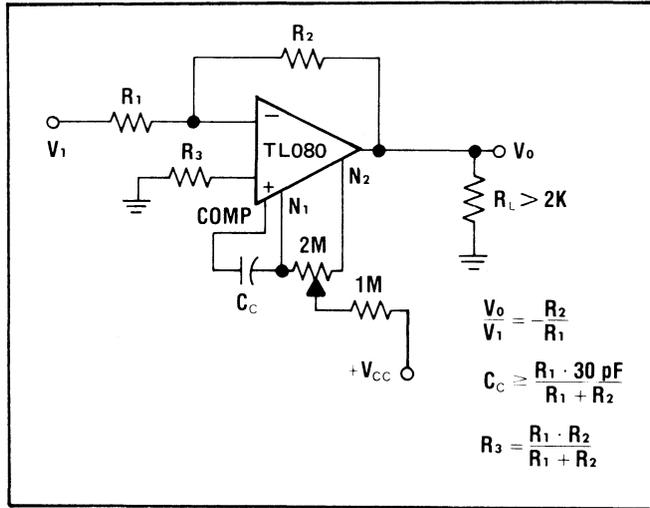


Fig. 17. — Inverseur à gain ajustable avec compensation de décalage d'entrée et de fréquence.

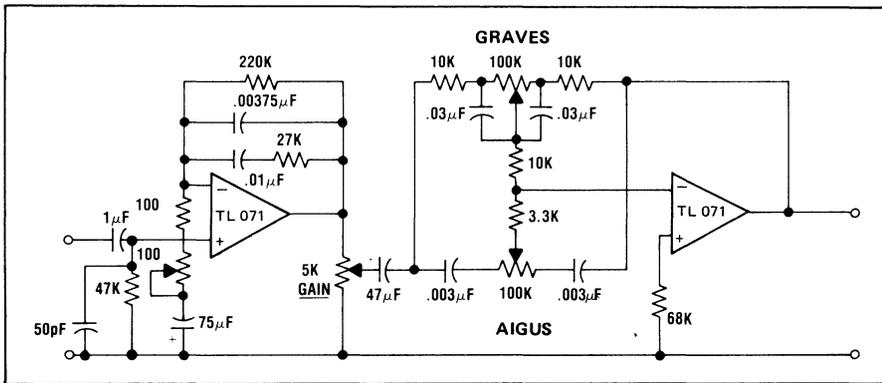


Fig. 18. — Préamplificateur audiofréquences.

La technologie BIFET

Dans le LF 156 de *National Semiconductor*, ce ne sont plus des MOS mais des FET à jonction qui sont utilisés, également dans une technologie mixte bipolaire-FET appelée « BIFET » (fig. 12). Il en résulte des performances typiques très remarquables, telles que :

- Courant de polarisation : 30 pA.
- Slew-rate : 15 V/µs.
- Produit gain-bande : 5 MHz.
- Offset à l'entrée : 3 mV.
- Dérive : 5µV/°C.
- Bruit : 12 n/√VHz.
- Temps d'établissement à 0,01 % : 1,5 µs.

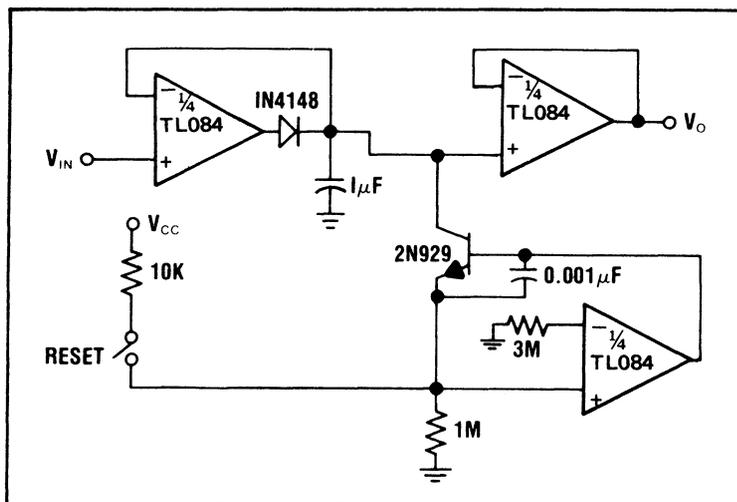


Fig. 19. — Détecteur de crête à haute impédance d'entrée, faible impédance de sortie et faible dérive.

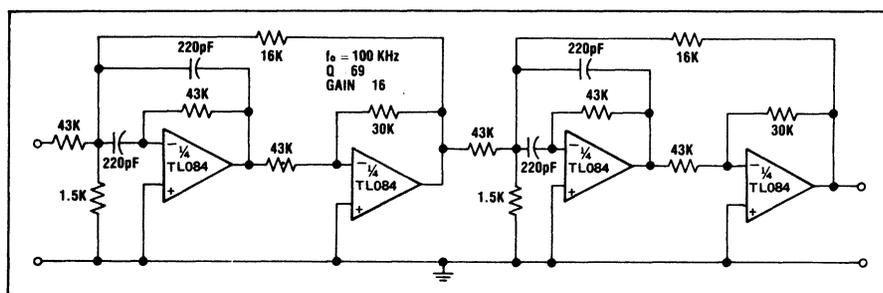


Fig. 20. — Filtre passe-bande du 2^e ordre. Avec $f_0 = 100$ kHz, on a $Q = 30$ et gain = 4; en cascade, on passe à $Q = 69$ et gain = 16.

Le LF 156 a servi de standard mais de nombreux fabricants ont, entre temps, proposé des familles qui leur sont propres. Parmi les plus performants, en 1978, on peut citer les HA-5105 de *Harris* et l'OP-16 de *Precision Monolithics Inc. (MMI)*. Le premier offre la bande passante remarquable de 15 MHz pour un slew-rate de $9 \text{ V}/\mu\text{s}$, un temps d'établissement de $1,7 \mu\text{s}$ à $0,01 \%$ et un temps de montée de 18 ns. Il est vrai qu'il est réalisé selon une technologie performante, l'isolation diélectrique. De plus, la finition du circuit est assurée par un ajustage dynamique au laser, avant encapsulation. Ainsi, l'offset d'entrée a été ramené à $0,5 \text{ mV}$ pour une dérive de $10 \mu\text{V}/^\circ\text{C}$.

L'OP-16, lui, se caractérise par une dérive excessivement faible, de $2 \mu\text{V}/^\circ\text{C}$, pour un élément de bande passante de 8 MHz et avec le slew-rate élevé de $25 \text{ V}/\mu\text{s}$. Pour obtenir cette faible dérive, avec un offset d'entrée de $0,5 \text{ mV}$, *PMI* a préféré réaliser un montage interne de compensation, avec « miroir de courant ». Son temps d'établissement est de 700 ns et son courant de bruit de $0,01 \text{ pA}$ par racine carrée de hertz.

Ces BIMOS et BIFET trouvent les mêmes domaines d'applications que les autres amplificateurs opérationnels : acquisition de données, amplification pour sortie de convertisseurs D/A, filtres actifs, amplification de photocellules, etc.

Quelques exemples de montages sont donnés dans les figures suivantes, proposées par *Texas* à partir de BIFET de sa production. On trouvera successivement un générateur de signaux carrés à 0,5 Hz (fig. 13), un circuit bouchon à coefficient de surtension élevé (14), un amplificateur de distribution audio (15), un oscillateur à quadrature avec fréquence centrale de 100 kHz (16), un inverseur à gain ajustable avec compensation d'offset et de fréquence (17), un préamplificateur audio (18), un détecteur de crête (19) et un filtre passe-bande (20); dans ce dernier cas, et avec montage en cascade, pour $F_o = 100$ kHz, on a $Q = 69$ et le gain = 16.

BIBLIOGRAPHIE

Peu d'ouvrages spécialisés ont été, à ce jour, consacrés exclusivement aux FET. Citons :

- [1] **MOS Integrated Circuits, Theory, Fabrication, Design, and Systems Applications of MOS LSI**, rédigé par l'équipe d'ingénieurs de **American Micro-Systems Inc [AMI]** en 1972, ouvrage le plus complet peut-être. Éditeur : Van Nostrand Reinhold, New York.
- [2] **MOSFET in Circuit Design**, par Robert H. CRAWFORD, publié dans l'excellente collection de *Texas Instruments* et datant de 1967. Distributeur en France : Société des Éditions Radio.
- [3] **Conception et technologie des circuits intégrés à transistors MOS**, par John MAJOR, éditions Eyrolles, résultant d'un séminaire organisé à Édimbourg en 1972 par l'*IEE*.
- [4] **COS/MOS integrated circuits manual**, édité par *RCA* et fondamental tant pour les MOS que pour les COS/MOS.
- [5] **Transistors à effet de champ**, par J. P. OEMICHEN, aux Éditions Radio, livre de base sur les JFET discrets.
- [6] **Field-Effect Transistors**, par L. J. SEVIN, dans la série des ouvrages de *Texas-Instruments*, où l'on trouvera développée la théorie des JFET.

En effet, les principales références bibliographiques (et leur liste est très longue) se rapportent aux analyses et études publiées dans les revues scientifiques sous les signatures de SCHOCKLEY, ARMSTRONG, HEIMANN, HOFSTEIN...

L'aspect pratique sera plus aisément appréhendé grâce aux documents publiés par les fabricants de CI/MOS, et en particulier par *General Instrument, Texas RCA, Motorola, Intel, National Semiconductor, SGS-Atés, Mostek, Fairchild, RTC, AMI* etc.

On pourra également se référer aux comptes rendus des grandes manifestations de microélectronique organisées par *EEE* et *EPN* à Paris (1970, 1971, 1972), les colloques de la *FNIE* sur les mémoires (Paris, 1973), les circuits intégrés complexes (Paris, décembre 1974), etc.

Enfin, la *Société des Éditions Radio* a édité plusieurs ouvrages de base consacrés à la microélectronique et publie, en particulier, la revue **Électronique et applications industrielles**, qui a consacré aux CI/MOS des études ponctuelles et de synthèse qui permettront de mettre et tenir à jour les connaissances sur ce sujet.

TABLE DES MATIÈRES

Liste des chapitres	5
Avant-propos	7

1^{re} PARTIE : FET ET CIRCUITS INTÉGRÉS

Chapitre I — Introduction aux FET	9
1. Les transistors unipolaires	9
2. Les avantages des FET	13
3. Les circuits à transfert de charge	14
4. SSI, MSI, LSI et VLSI	15
5. Fonctions logiques fondamentales	16
6. Caractéristiques principales des circuits intégrés	16
Temps de propagation et courbes caractéristiques	19
7. Les familles de circuits intégrés	19
8. L'histoire des FET	23

Chapitre II — Fonctionnement du MOS fondamental à enrichissement, canal P	27
1. Principe de fonctionnement	27
Action de la tension de drain-Zones triodes et saturée	30
2. Théorie de fonctionnement du MOS	33
La capacité de gate	33
La tension de seuil	34
Approximation de 1 ^{er} ordre	35
Le courant de drain	37
Région triode	37
Région pentode	37
Le rapport W/L et le facteur de gain β	38
Stabilité en température	39
La mobilité des porteurs P	39
La résistance drain-source	40
La transconductance	41
Théorie de 2 ^e ordre	42
Théorie de 3 ^e ordre	44
Le facteur de qualité	44
3. Quelques notions complémentaires	44
Le MOS parasite	44
La résistance d'entrée	45
Les capacités de structure	47
La protection des entrées	47
Autres limites en tension	49

Chapitre III — Les divers types de MOS	51
1. Introduction	51
2. Le MOS à canal N à enrichissement	52
Fonctionnement et caractéristiques	52
Les capacités de structure	56
Exemples de valeurs de courant drain	57
Les courants de fuite	59
3. Les MOS à déplétion	60
4. Le MOS tétrode (à double gate)	63
5. Le MOS à gate flottant	63
6. Le MNOS	66
7. Le MOS détecteur d'ions	66
Chapitre IV. — La fabrication des circuits intégrés	68
1. La technologie planar	68
Préparation du silicium	68
Oxydation	70
Photogravure	71
Dopage par diffusion gazeuse	74
Autres étapes	74
2. L'implantation ionique	75
Technique d'implantation des drain et source	77
Implantation de résistances	78
Implantation du canal	78
Fabrication d'inverseurs avec charge à déplétion	81
Fabrication de caissons pour MOS	81
L'usinage ionique	81
3. La présentation des CI	81
Boîtiers	81
Les puces nues	82
4. Les CI sur film	83
5. Techniques évoluées de masquage	85
Lithographie aux rayons X	87
Masquage au laser	87
Chapitre V. — L'arsenal des technologies	89
1. Introduction	89
2. Silicium 111	90
3. Silicium 100	91
4. Utilisation du nitrure de silicium	92
5. Gate au silicium	95
6. Le MOS à canal N	97
7. Le MOS à gate au molybdène (RMOS)	99
8. MOS à double diffusion (DMOS)	101
9. Les CMOS, ou MOS complémentaires	104
CMOS à gate au silicium	107
10. Circuits sur substrat isolant : le SOS	108
11. Techniques à oxyde dans le silicium	109
12. L'isolement diélectrique	110
Le thyristor parasite des CMOS	110
La technologie EPIC	112
13. MOS à couche mince	113
Les TFT du CNET	114
14. Le HMOS	117
15. La structure BMOS	118
16. Les BIMOS	118

Chapitre VI. — Les JFET et leurs familles	121
1. Le JFET de base	121
2. Le JFET Schottky	125
L'Astec de Thomson-CSF	127
Le TBM du LEP	129
3. Les Gridistors et le VFET	131
Les versions unipolaires : le VFET	131
Versions bipolaires ou mixtes	133
4. Les JFET en commutation	133
Le Varafet	138
5. Les BIFET	139

Chapitre VII. — Les circuits à transfert de charges : CCD, BBD et variantes 141

1. Principe des CCD	141
2. Les BBD	144
3. Les SCT	147
4. Les PCCD	147
5. Principes des applications	149
6. Registres et mémoires	149
Les disques en CCD	150
7. Les lignes à retard	150
8. Applications opto-électroniques	153
Méthodes d'analyse	153
L'amplificateur à gate flottant	154
Analyse à décharge de photodiode	155
Analyse linéaire	156
Analyse d'image	157
9. Le réseau à injection de charges	159
10. La logique à CCD	161

Chapitre VIII. — Les VMOS 167

1. Présentation des VMOS	167
2. VMOS en composants discrets	168
3. Circuits intégrés VMOS	169
4. Les VMOS discrets en commutation	171
5. Les VMOS discrets en linéaire	175

II^e PARTIE : LA LOGIQUE INTÉGRÉE A FET

Chapitre IX. — L'inverseur fondamental 179

1. L'inverseur à charge passive	179
La droite de charge	181
2. Le MOS utilisé comme charge	183
Principe	183
Fonctionnement dans la zone non saturée	183
Fonctionnement dans la zone saturée	185
Avantages et inconvénients	185
3. L'inverseur fondamental en commutation	186
Mise en service de l'inverseur	187
Blocage de l'inverseur	188

4. MOS à déplétion en charge	188
5. Inverseur à MOS complémentaires	191
Fonctionnement en commutation : la consommation	194
Le rôle de V_{DD} si l'on veut réduire la consommation	197
Immunité au bruit	201
6. Synthèse des inverseurs statiques	203
7. Le rôle de la capacité intrinsèque gate-substrat	203
8. La porte de transfert	205
9. Les inverseurs dynamiques	205
Inverseur proportionnel	205
Inverseur non proportionnel à précharge	206
Inverseur non proportionnel à précharge alimenté par l'horloge	209
Chapitre X. — Les portes logiques	211
1. Les portes et leurs caractéristiques	211
2. Portes statiques (monocanales)	213
Portes de base	213
OU exclusif	214
Autres portes	216
3. Le OU câblé et la sortie 3 états	216
4. Portes dynamiques	219
5. Portes CMOS	220
Porte de transfert	223
Porte NOR	223
Porte NAND	225
6. Interfaces pour MOS monocanaux	225
Bipolaire à MOS	227
MOS à bipolaire	229
Chapitre XI. — Les flip-flops	235
1. Les principales bascules	235
2. Le flip-flop RS	236
RS à inverseurs	238
RS à entrées multiples	238
RS synchrone	238
RS à sorties 3 états	239
RS en MOS complémentaires	240
3. Les RS dynamiques	241
4. La bascule JK	246
5. Le flip-flop T	251
6. Le flip-flop D	253
Le D à verrouillage	253
Le D à commutation sur les fronts	255
7. La bascule RST	255
8. Les bascules maître-esclave	257
9. Rôle des bornes « Preset » et « Clear »	257
10. Les flip-flops en CMOS	267
Les RS	258
Les flip-flops D	259
Chapitre XII. — Les registres	263
1. Principe des registres à décalage	263
2. Les types de registres	264
3. Les registres statiques biphasés	267
4. Les registres dynamiques biphasés à inverseurs proportionnels	267
5. Registre dynamique proportionnel triphase	270

6. Registres dynamiques non proportionnels à précharge	273
7. Registres dynamiques alimentés par les horloges	277
8. Registres en CMOS	280
9. Registres à longueur ajustable	280
10. Registres rebouclés	281
Chapitre XIII. — Les MOS complémentaires (CMOS)	283
1. Une introduction aux CMOS	283
2. La porte CMOS	285
Remarque à propos de la consommation	288
La mise en parallèle des entrées	289
L'immunité au bruit	289
Le OU câblé et la logique 3 états	293
Logiques positives et négatives	294
3. L'interfaçage des CMOS avec les autres logiques	295
TTL à CMOS	295
CMOS à TTL	295
Le décalage des niveaux	297
CMOS et HTL	297
CMOS à ECL	299
ECL à CMOS	300
CMOS et MOS monocanaux	301
CMOS à systèmes industriels et à d'autres composants	302
4. Les impulsions d'horloge	305
5. L'alimentation des CMOS	305
6. Les CMOS/SOS et le fonctionnement à forte déplétion	308
Le fonctionnement à forte déplétion	310
Les niveaux d'interconnexion	311
Quelques caractéristiques influant sur la vitesse	312
La variante ESFI de Siemens	313
La résistance aux radiations ionisantes	314
7. La 54C/74 C de NS	315
8. La 34000 Isoplanar de Fairchild	315
9. L'isolement diélectrique de Harris Semiconductor	316
10. Le substrat en spinelle	316
11. CMOS ou I ² L?	317
12. Logique majoritaire et CMOS	318

III° PARTIE : LES APPLICATIONS

Chapitre XIV. — Les mémoires mortes	327
1. Classement des mémoires	327
2. Organisation d'une mémoire	329
3. Caractéristiques	330
4. Principe des ROM	331
5. Les PROM	333
PROM à fusible	334
PROM à court-circuit de jonction	336
6. EPROM ET REEPROM	337
EPROM à effacement par ultraviolet	337
7. EAROM et EEROM	338
Structures de EFCIS en SOS	340
8. Les réseaux logiques programmables	341

Chapitre XV. — Les mémoires vives	345
1. La prodigieuse évolution des cellules RAM	345
2. RAM statiques à MOS monocanaux	347
Cellule à 8 transistors	347
Cellule à 6 transistors : la 1101	348
Cellule à charges à déplétion : la 2102	349
3. RAM statiques en CMOS	350
4. La volatilité des RAM	352
5. RAM pseudo-dynamiques	353
6. RAM dynamiques	355
7. L'avenir des RAM	357
Chapitre XVI. — Les microprocesseurs	359
1. Logique câblée et logique programmée	359
2. Structure d'un microprocesseur	360
3. Exemple d'un microprocesseur, le 8085	363
4. Exemple d'un micro-ordinateur d'enseignement	366
5. Les trois catégories de microprocesseurs	367
6. Le logiciel	372
Chapitre XVII. — Télécommunications : les CODEC	379
1. Principes	379
La compression expansion	380
2. Comment fonctionne un CODEC	382
3. L'exemple de CODEC types	384
Monovoies et multivoies	385
4. Marchés et applications	390
Chapitre XVIII. — Applications linéaires	393
1. Le MOS en linéaire	393
2. Applications AF	394
Amplificateur BF	394
MOS avec émetteur-suiveur	395
3. Structure différentielle	395
4. Montages à CMOS	397
5. La boucle à phase asservie	397
6. Les amplificateurs opérationnels	398
L'amplificateur à chopper	398
L'amplificateur BIMOS	401
La technologie BIFET	405
Bibliographie	409

au service de l'électronique d'aujourd'hui

1. électronique professionnelle

électronique
actualités

Journal hebdomadaire des cadres de l'industrie électronique

Electronique
applications industrielles

Revue technique bimensuelle d'applications industrielles de l'électronique

Toute l'Electronique

Revue mensuelle des ingénieurs et techniciens de l'électronique

automatique
& informatique
industrielles

Revue mensuelle des techniques et applications industrielles de l'automatisation

2. électronique grand public

électronique pour vous



Magazine

Le magazine qui fait autorité par la valeur de ses études.

la nouvelle
REVUE DU SON

DES IDEES ■ DES NOUVEAUTES ■ TOUS LES PRIX

La revue du véritable audiophile.



9, rue Jacob - 75006 PARIS
Tél. 033.13.65

EXTRAIT DU CATALOGUE DES EDITIONS RADIO

L'ORDINATEUR ET L'INFORMATIQUE en 15 leçons, par P. Morvan — Les connaissances nécessaires et suffisantes pour comprendre l'informatique.
224 pages, format 14 x 20

COURS FONDAMENTAL DE LOGIQUE ELECTRONIQUE, par R. Amato — Avec des exemples, l'auteur explique le langage et le fonctionnement des éléments de base en logique électronique. Puis il conduit le lecteur à des réalisations complexes. Une excellente introduction à l'ordinateur.
328 pages, format 16 x 24 (2ème édition)

CIRCUITS DE LOGIQUE, par R. Damaye — Circuits élémentaires, notion de comptage électronique, réalisation et protection, algèbre logique.
372 pages, format 16 x 24 (4ème édition)

LOGIQUE ELECTRONIQUE ET CIRCUITS INTEGRES NUMERIQUES, par R. Damaye — Technologie, principe de fonctionnement électrique et logique, schémas d'utilisation des circuits intégrés de logique.
528 pages, format 16 x 24 (3ème édition)

EMPLOI RATIONNEL DES CIRCUITS INTEGRES, par J.-P. Oehmichen — Structures, circuits logiques, mémoires, comptage binaire et décimal, emploi des circuits intégrés, amplificateurs opérationnels, comparateurs, etc.
512 pages, format 16 x 24 (3ème édition)

DU MICROPROCESSEUR AU MICRO-ORDINATEUR, par H. Lilen — Une introduction à la micro-informatique de la théorie à la pratique.
448 pages, format 16 x 24 (3ème édition revue et corrigée)

THEORIE ET PRATIQUE DES MICROPROCESSEURS, par R. Arouète et H. Lilen — Etude fondamentale des logiques programmées et des microprocesseurs (matériels, logiciels, mise en œuvre, coûts).
160 pages, format 16 x 24 (2ème édition)

COMPRENDRE LES MICROPROCESSEURS EN 15 LEÇONS, par D. Queyssac — Pour mettre à jour ses connaissances sur les microprocesseurs, clés de l'électronique actuelle et future.
160 pages, format 14 x 20

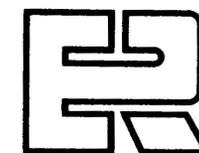
GUIDE MONDIAL DES MICROPROCESSEURS, par H. Lilen — Les critères de sélection, les caractéristiques et le classement des microprocesseurs.
192 pages, format 21 x 29,7

CIRCUITS INTEGRES NUMERIQUES, par H. Lilen — Principes et applications des circuits intégrés numériques : familles, fonctions, utilisation.
384 pages, format 16 x 24 (3ème édition)

MEMOIRES INTEGRES, par H. Lilen — Fonctionnement, choix, utilisation.
288 pages, format 16 x 24

PROGRAMMATION DES MICROPROCESSEURS, par H. Lilen — Acquerez avec ce livre la maîtrise des logiques programmables, clé de l'immense domaine des applications des microprocesseurs.
240 pages, format 16 x 24 (3ème édition)

ENVOI DE NOTRE CATALOGUE SUR DEMANDE



Editions Radio

9, RUE JACOB - 75006 PARIS